

**Quasilösung
von
Parameteridentifikationsproblemen
im
Banachraum**

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

Doctor rerum naturalium

von

Andrej Klassen

bei der Fakultät für Mathematik der Universität Duisburg-Essen vorgelegt

Gutachter der Dissertation:

Prof. Dr. Christian Clason und *Prof. Dr. Barbara Kaltenbacher*

Tag der mündlichen Prüfung:

21.10.2019

DANKSAGUNG

An dieser Stelle möchte ich meinen Dank der Universität Duisburg-Essen für das Promotionsstipendium entgegenbringen und den Menschen, ohne die es diese Arbeit so nicht gegeben hätte.

Als erstes danke ich meinem Betreuer Prof. Dr. Christian Clason, der für meine Anliegen und Ideen immer ein offenes Ohr hatte und durch kritische Fragen und Anregungen mich auf den rechten Weg gebracht hat, und insbesondere dafür, dass er mich nicht nur von der fachlichen, sondern auch von der menschlichen Seite immer unterstützte. Ich danke auch Prof. Dr. Barbara Kaltenbacher für sehr interessante und fruchtbare Zusammenarbeit, für viele hilfreiche Ratschläge und Feedbacks, für ihre Wertschätzung und Hilfsbereitschaft. Des Weiteren danke ich meinem Diplomvater Prof. Dr. Arnd Rösch, denn ohne ihn wäre die Möglichkeit zu dieser Promotion nicht entstanden. Schließlich danke ich meiner Familie für den permanenten Rückhalt.

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub | universitäts
bibliothek

Diese Dissertation wird über DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt und liegt auch als Print-Version vor.

DOI: 10.17185/duepublico/70737

URN: urn:nbn:de:hbz:464-20191125-101335-1

Alle Rechte vorbehalten.

INHALTSVERZEICHNIS

1	EINLEITUNG	5
2	GRUNDLAGEN UND ZIELSETZUNG	7
2.1	Allgemeines Regularisierungsschema	7
2.2	Ivanovs Quasilösungen	11
2.3	Schwache Kompaktheit in Banachräumen	14
2.4	Zielsetzung	17
3	QUASILÖSUNGEN BEI LINEAREN IDENTIFIKATIONSPROBLEMEN	19
3.1	Existenz der Quasilösungen	19
3.2	Regularisierung	22
3.3	Distanzfunktion	25
3.4	Parameterwahl	29
3.5	Konvergenzraten	33
3.6	Anwendung: Numerische Lösung, linearer Fall	37
4	QUASILÖSUNGEN BEI NICHTLINEAREN IDENTIFIKATIONSPROBLEMEN	45
4.1	Existenz und Stabilität	45
4.2	Zur Ermittlung des Regularisierungsparameters	49
4.3	Konvergenzraten	53
4.4	Anwendung: Numerische Lösungen, semi-linearer Fall	55
5	ZUSAMMENFASSUNG UND AUSBLICK	61
	Literatur	62

1 EINLEITUNG

Als Parameteridentifikationsproblem wird in der Naturwissenschaft typischerweise die Interpretation indirekter Messungen bezeichnet. Bei dieser Interpretation sollen zu den Daten, welche durch indirekte Messungen gegeben sind, gewisse physikalische Größen (Parameter) gefunden werden. Zur Verfügung steht ein, durch eine Operatorgleichung beschriebenes, mathematisches Modell der direkten Aufgabe, welches den Zusammenhang zwischen Daten und den ursprünglichen Parametern beschreibt. Dieses Ursache–Wirkung Modell weist in der Regel stark glättende Eigenschaften auf, so dass das zu lösende Inverse Problem, neben der oft auftretenden Mehrdeutigkeit einer Lösung zusätzlich Instabilität in sich birgt. Der Umgang mit der Instabilität zählt zu den schwierigsten Hürden bei der Lösung von Parameteridentifikationsproblemen. Darauf entstand Mitte des 20. Jahrhunderts eine grundlegende Idee, solche Probleme auf Basis von Regularisierungen approximativ zu lösen. Dabei besteht das Prinzip einer Regularisierung darin, nicht das exakte instabile Ausgangsproblem zu lösen, sondern ein naheliegendes stabiles. Die Abweichung wird also zugunsten der Stabilität bewusst in Kauf genommen und die gemessenen Daten werden approximiert - "nicht so genau wie möglich, sondern nur so genau wie nötig" [29].

In den letzten 70 Jahren ist in der Mathematik eine Reihe von Methoden zur Analyse und approximativem Lösen solcher Probleme entwickelt worden [53], [33], [43]. Eine der bekanntesten, welche sich gegenüber den anderen stark durchgesetzt hat, ist sicherlich die von Tichonov vorgeschlagene Vorgehensweise. Dabei wird der Operator der direkten Gleichung durch einen solchen approximiert, der eine stetige Inverse besitzt [53], [52]. Gekoppelt an der sich rasant entwickelnden numerischen Mathematik, hat dieser Ansatz einen großen Erfolg in der Theorie und auch in der Praxis gewährleistet. Dieser Erfolg ist in einigen angewandten Disziplinen, welche erst mit den neuen Methoden einen mathematischen Framework zur Lösung bekommen haben, wie z.B. mathematische Bildverarbeitung, besonders gut zu sehen [49], [7].

Am Anfang wurden die Regularisierungsmethoden auf Hilberträumen definiert. Im Laufe der Entwicklung wurden diese Forderungen abgeschwächt. Jedoch gibt es momentan nicht viele Veröffentlichungen, welche sich mit der Möglichkeit einer Regularisierung auf nicht-reflexiven Banachräumen befassen [11], [8]. Die Abschwächung der Forderung an die Funktionenräume veranlasst dazu ein breiteres Spektrum an Regularisierungsansätzen in Betracht zu ziehen. Ein solcher, mehr oder weniger in den Schatten der Tichonov Methode geratener Ansatz, basiert auf der Einschränkung des Definitionsbereiches und geht auf einen bekannten Satz zurück, welcher besagt, dass für eine stetige Abbildung mit kompaktem Definitionsbereich die Inverse Abbildung stetig ist. Dieser Satz wird zum Fundament einer Regularisierungsideologie, welche von V. K. Ivanov vorgeschlagen wurde, und zwar die Methode der Quasilösungen und der expandierenden Kompakte [33], [34]. Daher wird sie in der Literatur auch als die Ivanov–Regularisierung bezeichnet. Allerdings gibt es oft große Schwierigkeiten durch be-

stimmte Einschränkungen eine kompakte Menge in unendlich-dimensionalen Räumen zu konstruieren, so dass Varianten mit der Abschwächung der Forderung zum Definitionsbereich (z.B. schwach-kompakte Lösungsmengen) untersucht wurden [19], [35], [36].

Das Ziel dieser Arbeit ist eine theoretische Analyse und Begründung der Anwendbarkeit der Ivanov-Regularisierung für Parameteridentifikationsprobleme im Fall eines nicht-reflexiven Lösungsraumes. Dabei wird vor allem eine Art der abgeschwächten Regularisierungseigenschaften für schwach*-kompakte Mengen als Definitionsbereich untersucht, optimale Regel zur Parameterwahl vorgeschlagen und die Konvergenzeigenschaften der Lösungen analysiert. Darunter wird auch die numerische Umsetzung für exemplarische Testprobleme implementiert. Als Lösungsraum wird der Banachraum L^∞ gewählt. Grund dafür, abgesehen von der praktischen Bedeutsamkeit, sind die "schlechten Eigenschaften" dieses Raumes: nicht-reflexiv, nicht-separabel, nicht-strikt-konvex. Also solche Eigenschaften, welche alle gängigen Regularisierungsmethoden vor große Schwierigkeiten stellen. Einige Resultate dieser Arbeit sind in [14], [38], [39] publiziert worden, wobei bestimmte wichtige Aspekte der letzteren zwei Publikationen über den Umfang dieser Arbeit gehen. Dazu gehören unter anderem die Konvergenzanalyse der Morosov-Regularisierung [38] und der Aufbau und Untersuchung eines iterativen Vorgehens auf Basis von Ivanovs Quasilösungen [39].

Die Arbeit ist wie folgt gegliedert:

Im Kapitel 2 werden einige Ideen und die bedeutendsten Definitionen und Sätze, welche die Grundlagen zur theoretischen Analyse in dieser Arbeit darstellen, kurz vorgestellt und diskutiert. Das Kapitel endet mit einer ausführlichen Aufgabenstellung der vorliegenden Arbeit.

Im Kapitel 3 wird der Einsatz der Ivanov-Regularisierung mit vorausgesetzten Forderungen an den Lösungsraum für Lineare Probleme begründet. Dabei wird das Verhalten der Quasilösungen in Anlehnung an die klassische Vorgehensweise bei der Prüfung von Regularisierungseigenschaften analysiert, gängige Parameterwahlstrategien untersucht und Konvergenzraten ermittelt. Anschließend werden die Algorithmen und praktische Realisierung auf Basis von Semi-Smooth-Newton Methoden für den linearen Fall dargestellt.

Im Kapitel 4 wird die prinzipielle Anwendbarkeit des Verfahrens für den Fall der nichtlinearen Operatorgleichungen gezeigt. Das Kapitel endet ebenfalls mit einer numerischen Lösung für ein semi-lineares Testproblem.

2 GRUNDLAGEN UND ZIELSETZUNG

2.1 ALLGEMEINES REGULARISIERUNGSSCHEMA

Historisch gesehen entstand eine Regularisierung als mathematisches Verfahren zur Behandlung inkorrekt linearer Operatorgleichungen im Hilbertraum. Obwohl eine Regularisierung im Hilbertraum nicht zum Schwerpunkt dieser Arbeit gehört, wird in diesem Abschnitt ein kurzer Überblick über dieses Thema gemacht, um auf die gängigen Begriffe, wichtige Definitionen und einige Resultate einzugehen, welche für die weitere Darstellung von Bedeutung sind.

Für eine lineare und stetige Abbildung $A : X \rightarrow Y$ auf den Hilberträumen X und Y heißt die Operatorgleichung

$$(2.1.1) \quad Ax = y$$

korrekt, wenn eine Lösung x zu jedem y existiert, eindeutig ist, und stetig von der rechten Seite y abhängt [50, Def. 3.1]. Diese Definition eines *korrekt gestellten Problems* geht auf *Hadamard* zurück. Die letzte Forderung wird in der Literatur als *Stabilität* bezeichnet.

Die *Moore-Penrose Inverse* A^\dagger zum Operator A wird als die eindeutige lineare Erweiterung von \tilde{A}^{-1} definiert, so dass der Definitionsbereich $D(A^\dagger) = R(A) \oplus R(A)^\perp$ mit $N(A^\dagger) = R(A)^\perp$, wo $\tilde{A} = A|_{N(A)^\perp} : N(A)^\perp \rightarrow R(A)$ [20, Def. 2.2] ist. Damit ist \tilde{A} wohldefiniert und injektiv auf $N(A)^\perp$ und besitzt eine Inverse. Der Operator A^\dagger ist also eine verallgemeinerte Inverse des linearen Operators A im Hilbertraum, so dass für $y = y_1 + y_2$ mit $y_1 \in R(A)$, $y_2 \in R(A)^\perp$ gilt $A^\dagger y = \tilde{A}^{-1}y_1$.

Die Moore-Penrose Inverse zum Operator A ist genau dann stetig, wenn $R(A) \subset Y$ abgeschlossen ist [20, Prop. 2.4], also genau dann, wenn die Operatorgleichung (2.1.1) *korrekt gestellt nach Nashed* ist [50, Def. 3.5].

Bei der Behandlung von Parameteridentifikationsproblemen muss damit gerechnet werden, dass in der Realität nur eine Näherung der rechten Seite y bekannt ist. Man spricht daher über die *exakte Daten* $y \in Y$ und *gestörte Daten* $y^\delta \in Y$. Dabei wird die Störung mit einem als bekannt vorausgesetzten *Störungslevel* $0 < \delta < \|y^\delta\|$ wie folgt definiert

$$\|y - y^\delta\| \leq \delta$$

Es wird angenommen, dass die exakten Daten $y \in Y$ im Definitionsbereich der Moore-Penrose Inverse liegen, dann ist $x^\dagger = A^\dagger y$ eine Lösung der Gauss-Normalgleichung

$$(2.1.2) \quad A^* Ax = A^* y$$

mit der minimalen Norm. Doch sogar im unwahrscheinlichen Fall $y^\delta \in D(A^\dagger)$, stellt $A^\dagger y^\delta$ keine gute Approximation von x^\dagger dar, wenn die Operatorgleichung (2.1.1) inkorrekt nach Nashed ist, da die verallgemeinerte Inverse unstetig ist.

Zur Konstruktion der approximativen Lösung von inkorrekten Operatorgleichungen ist in der Mathematik eine Technik mit dem Oberbegriff *Regularisierung* entwickelt worden. Dabei wird ein inkorrektes Problem durch eine naheliegende parametrisierte Familie von korrekten Problemen approximiert, so dass sich ein klassisches Regularisierungsschema herauskristallisiert hat vgl. [20], [50], [29]. Zur Konstruktion einer approximativen Lösung der Gleichung (2.1.1) mit der gestörten rechten Seite y^δ wird eine *Parametrisierte Familie von Operatoren* $\{R_\alpha\}$ mit dem *Regularisierungsparameter* $\alpha \in (0, \infty)$ betrachtet.

Definition 2.1. Eine Familie $\{R_\alpha\}_{\alpha>0}$ heißt *Regularisierung für den Operator* A^\dagger , wenn

- (i) $R_\alpha y^\delta \rightarrow R_\alpha(y)$, wenn $\delta \rightarrow 0$, für jedes $\alpha > 0$, jedes $y \in Y$
- (ii) $R_\alpha y \rightarrow A^\dagger(y)$, wenn $\alpha \rightarrow 0$, für alle $y \in R(A) \oplus R(A)^\perp$
- (iii) Es existiert eine Regel zur Parameterwahl

$$(\delta, y^\delta) \mapsto \alpha(\delta, y^\delta) \quad \text{so, dass für } \alpha(\delta, y^\delta) \rightarrow 0 \text{ gilt}$$

$$\limsup_{\delta \rightarrow 0} \left\{ \|R_\alpha y^\delta - x^\dagger\|_X : \|y - y^\delta\|_Y < \delta \right\} = 0$$

Ist $\{R_\alpha\}_{\alpha>0}$ eine Regularisierung von A^\dagger für inkorrekte Operatorgleichung mit

$$\sup \{ \|AR_\alpha\| : \alpha > 0 \} < \infty$$

so gilt für alle $y \in Y \setminus D(A^\dagger)$, dass $\|R_\alpha y\| \rightarrow \infty$ bei $\alpha \rightarrow 0$ s. [20, Prop. 3.6]

Man unterscheidet drei Strategien zur Wahl des Regularisierungsparameters

Definition 2.2. Die Funktion $\alpha : \mathbb{R}^+ \times Y \rightarrow \mathbb{R}^+$ definiert eine *Parameterwahlstrategie*. Dabei unterscheidet man die Strategien wie folgt

- (i) *a-priori* $\delta \mapsto \alpha(\delta)$
- (ii) *heuristische* $y^\delta \mapsto \alpha(y^\delta)$
- (iii) *a-posteriori* $(\delta, y^\delta) \mapsto \alpha(\delta, y^\delta)$

Im Bezug auf die klassische Definition einer Regularisierung ist eine wichtige Tatsache wohl bekannt s. z.B. [20, Prop. 3.4]

- Falls für eine Operatorfamilie $\{R_\alpha\}_{\alpha>0}$ die Forderungen (i) und (ii) der Definition 2.1 gelten, dann existiert stets eine a-priori Parameterwahl $\alpha = \alpha(\delta)$, so dass die Familie $\{R_\alpha\}$ eine Regularisierung bildet, d. h. die Forderungen (iii) der Definition 2.1 ist dann ebenfalls erfüllt.

Dieses Ergebnis macht die dritte Forderung der Definition 2.1 sicherlich nicht überflüssig, da die parametrisierte Operatorfamilie immer nur im Zusammenhang mit einer Parameterwahlregel ein Regularisierungsverfahren bildet. Auch die anderen Parameterwahlstrategien sind von großer Bedeutung. Die optimale Wahl von $\alpha = \alpha(\delta)$ benötigt zum Beispiel oft

zusätzliche Informationen über die Lösung x^\dagger , welche bei der praktischen Umsetzung kaum vorhanden sind. Bei einer a-posteriori Parameterwahl $\alpha = \alpha(\delta, y^\delta)$ wird die Information über die gestörten Daten mit einbezogen, so dass man ohne zusätzliche Informationen auskommt. Eine heuristische Strategie $\alpha = \alpha(y^\delta)$ verzichtet dagegen auf die Kenntnis des Fehlerniveaus.

Der oben beschriebenen klassischen Regularisierungstheorie liegt größtenteils die von Tichonov vorgeschlagene Vorgehensweise zugrunde [53], [52]. Dabei wird die inkorrekte Gleichung 2.1.1 durch eine Familie von Nachbargaufgaben

$$(2.1.3) \quad (A^*A + \alpha I)x = A^*y$$

ersetzt. Die Familie der regularisierenden Operatoren ist dann durch den Parameter α wie folgt gegeben

$$R_\alpha = (A^*A + \alpha I)^{-1}A^*$$

Diese Regularisierung wird oft in einer equivalenten, *variationellen Formulierung*, als ein Optimierungsproblem

$$T_\alpha(x) := \frac{1}{2} \|Ax - y^\delta\|_Y^2 + \|x\|_X^2 \rightarrow \min \quad x \in X$$

formuliert und $T_\alpha : X \rightarrow \mathbb{R}$ als das *Tichonov-Funktional* bezeichnet.

Es ist leicht zu prüfen, dass die Tichonov-Regularisierung bei einer a-priori Parameterwahl $\alpha = \alpha(\delta) = c\sqrt{\delta}$ alle Kriterien der Definition 2.1 erfüllt. Dabei können die *Konvergenzraten* von

$$\|R_\alpha y^\delta - x^\dagger\| = O(\sqrt{\delta})$$

erreicht werden s. z.B. [20, Thm. 10.4], wenn ein Element $v \in Y$ existiert, so dass die exakte Lösung durch

$$x^\dagger = A^*v$$

darstellbar ist. Solche Informationen über die exakte Lösung werden als *Quelldarstellungen* bezeichnet. Eine nicht zu vernachlässigende Eigenschaft einer Regularisierung stellt die *Qualifikation* eines Verfahrens dar s. [48, Def. 3.3.8], die zusammen mit entsprechenden Quelldarstellungen die optimale Parameterwahl ermöglicht. So können z.B. bei vorhandener Quelldarstellung

$$\exists w \in X : \quad x^\dagger = A^*Aw$$

und einer a-priori Parameterwahl $\alpha = \alpha(\delta) = c\delta^{2/3}$ die Konvergenzraten von

$$\|R_\alpha y^\delta - x^\dagger\| = O(\delta^{2/3})$$

erreicht werden. Dieses Resultat kann bei einem iterativen Vorgehen in einer Regularisierung noch besser ausfallen s. [57].

Die Frage der Regularisierbarkeit (also die Existenz von mindestens einem Regularisierungsalgorithmus) zu der Gleichung (2.1.1) ist bei einer Lockerung der Forderungen an die Funktionenräume nicht mehr so einfach zu beantworten. Sogar im Fall eines linearen und stetigen Operators $A : X \rightarrow Y$ müsste diese Frage, für den Fall X und Y seien lineare, normierte

2 Grundlagen und Zielsetzung

Räume, negativ beantwortet werden. So zeigte V.A. Vinokurov zur notwendigen Bedingung der Tichonov-Regularisierbarkeit, dass die lineare Operatorgleichung

$$Ax = y$$

auf Banachräumen X und Y nicht regularisierbar ist, wenn der Datenraum Y separabel, und der Lösungsraum X nicht-separabel ist [60]. Dabei wurde die Regularisierbarkeit im klassischen Sinne s. Definition 2.1 verstanden, also mit der erwarteten starken Konvergenz der regularisierten Lösungen gegen exakte für $\delta \rightarrow 0$.

2.2 IVANOV'S QUASILÖSUNGEN

Neben der Tichonov-Regularisierung stellen die Ivanovs Quasilösungen ein mächtiges Werkzeug zur Behandlung inkorrekt gestellter Operatorgleichungen dar, welches auch in der Gegenwart an Bedeutung nicht verloren hat s. [59], [63].

Ursprünglich wurde die *Methode der Quasilösungen* in [33]; [34] für lineare Gleichungen in normierten Räumen formuliert. Die Idee der Methode basiert auf einem bekannten topologischen Satz s. [54].

Satz 2.3. *Sei die Abbildung $A : X \rightarrow Y$ auf normierten Räumen X, Y injektiv und stetig. Sei weiterhin $M \subset D(A) \subset X$ eine kompakte Menge. Dann ist die inverse Abbildung $A^{-1} : A(M) \rightarrow M \subset X$ stetig.*

Wenn alle möglichen gestörte Daten y^δ zu der Bildmenge $A(M)$ gehören würden, so könnte nach dem Satz 2.3 die inkorrekte Operatorgleichung (2.1.1) als *bedingt korrekt* auf dem Kompakt $M \subset D(A)$, in der Literatur auch als *korrekt nach Tichonov* bezeichnet, betrachtet werden. Das heißt, mit einer Einschränkung des Definitionsbereiches des Operators könnte man theoretisch die bedingte Korrektheit zurückgewinnen. Die Erfüllung der Forderung, dass alle y^δ dabei zum Bild der kompakten Menge gehören, ist in unendlich-dimensionalen Räumen aber kaum zu erwarten, so dass die Existenz der Lösungen zu (2.1.1) nicht gewährleistet ist. Daraufhin wurde 1962 von V.K. Ivanov die *Methode der Quasilösungen* vorgeschlagen [33]; [34].

Definition 2.4. *Für eine kompakte Menge $M \subset D(A) \subset X$ wird ein Element $\tilde{x} \in M$ zu einer gegebenen rechten Seite $y^\delta \in Y$ als die Quasilösung der Operatorgleichung (2.1.1) bezeichnet, wenn gilt*

$$(2.2.1) \quad \|A\tilde{x} - y^\delta\| = \inf \{ \|Ax - y^\delta\|_Y : x \in M \}$$

Wenn M zusätzlich konvex und die Räume X und Y gleichmäßig konvexe Banachräume sind, existiert die eindeutige Quasilösung [19] und es gilt

$$(2.2.2) \quad \tilde{x} = \arg \min \{ \|Ax - y^\delta\|_Y : x \in M \}$$

Darüber hinaus ist bei obigen Voraussetzungen zu den Funktionenräumen die *metrische Projektion*

$$\tilde{y} \mapsto P_{A(M)}(y^\delta)$$

gegeben durch die Gleichung (2.2.1), stetig. Dabei definiert

$$P : Y \rightarrow A(M) \subset Y$$

den *Projektionsoperator* in Y . So ist aber nach Satz 2.3 auch $A^{-1} \circ P$, als eine Komposition stetiger Operatoren ebenfalls stetig, so dass die Quasilösungen *stabil* von den Eingangsdaten y^δ abhängen. Das heißt für eine Datenfolge $y^{\delta_n} \rightarrow y$ bei $\delta_n \rightarrow 0$ gilt für die entsprechenden Quasilösungen $\{\tilde{x}_n\}$

$$\lim_{n \rightarrow \infty} \tilde{x}_n = \arg \min \{ \|Ax - y\|_Y : x \in M \}$$

2 Grundlagen und Zielsetzung

Wenn sich eine parametrisierte Familie $\{M_{\rho_n}\}_{n \in \mathbb{N}}$ von eingebetteten kompakten Mengen $M_{\rho_n} \subset M_{\rho_{n+1}}$ für $\rho_n < \rho_{n+1}$ konstruieren lässt (später in der Literatur als die *Methode der expandierenden Kompakte* bezeichnet, bzw. *method of extending compacts* s. [63]), so bieten die Quasilösungen ein Regularisierungsverfahren [19], [35], [36].

Allerdings gibt es gewisse Schwierigkeiten in der Konstruktion von kompakten Mengen in unendlich-dimensionalen Funktionenräumen. Dies stellt eine natürliche Einschränkung für einen breiten praktischen Einsatz dieses Verfahrens dar. Hilfreich sind dabei oft die a-priori gegebenen Informationen über die exakte Lösung x^\dagger von (2.1.1). So wird z.B. eine Art der Quelldarstellung genutzt, welche bei einigen physikalischen Problemen gegeben ist: - für einen kompakten linearen Operator $L : Z \rightarrow X$, welcher auf dem Hilbertraum Z definiert ist, ist das Bild einer Kugel $B_\rho \subset Z$ mit dem Radius ρ kompakt in X . So dass folgender Algorithmus, bestehend aus drei Schritten, aufgebaut werden kann s. [63]

- (i) wähle ρ_1 als Anfangswert für $n = 1$, und eine Schrittweite $d_\rho > 0$
- (ii) finde die entsprechende Quasilösung \tilde{x}_n von (2.2.1) auf der Menge $M = L(B_{\rho_n})$
- (iii) wenn $\|A\tilde{x}_n - y^\delta\| > \delta$ setze $\rho_{n+1} = \rho_n + d_\rho$ und gehe zu (ii), sonst ist \tilde{x}_n die Lösung

In den späteren Veröffentlichungen [19] und [37] ist der Zusammenhang zwischen der Methode der Quasilösungen und der Tichonov-Regularisierung untersucht worden.

Satz 2.5. *Seien X und Y gleichmäßig konvexe Banachräume, die linearen Abbildung $A : X \rightarrow Y$ stetig und $L : X \rightarrow X$ kompakt. Sei $R(A) \equiv Y$, d.h. das Bild des Operators A ist dicht in Y . Sei weiterhin ein $y \in Y$ beliebig aber fest gegeben. Dann gilt*

- (i) für jedes $\rho_0 \in [0, \infty)$ existiert eine eindeutige Lösung x_0 des Minimalproblems

$$\|Ax - y\|^p \rightarrow \min \quad \text{für } x \in \{x \in X : \|Lx\|^q \leq \rho_0^q\}$$

- (ii) es existiert ein $\alpha_0 > 0$, so dass x_0 die eindeutige Lösung des Minimalproblems

$$\|Ax - y\|^p + \alpha \|Lx\|^q \rightarrow \min \quad x \in X$$

ist mit $\alpha = \alpha_0$.

Beweis s. [19] und [37, Thm. 5;Kap. 3]. In [45] wurde dieses Resultat für einen nicht-linearen Operator $F : X \rightarrow Y$ unter Voraussetzung der Konvexität von $\|F(x) - y^\delta\|^2$ festgestellt und ein Algorithmus zur Ermittlung von α_0 vorgeschlagen.

Das Thema des Zusammenhangs zwischen verschiedenen variationellen Regularisierungsvorgehensweisen wurde auch in [44], [51] angesprochen. In [41] werden drei, in der Literatur wohl bekanntesten Verfahren, in der verallgemeinerten Form dargestellt

- Tichonov-Regularisierung

$$(T) \quad \min_{x \in X} G(F(x), y) + \alpha J(x)$$

- Morozov-Regularisierung (oder Residualmethode)

$$(M) \quad \min_{x \in X} J(x) \quad \text{s.t.} \quad G(F(x), y) \leq \delta$$

- Ivanov-Regularisierung (oder Quasilösungsmethode)

$$(I) \quad \min_{x \in X} G(F(x), y) \quad \text{s.t.} \quad J(x) \leq \rho$$

betrachtet mit einem stetigen, aber nicht notwendigerweise linearen Operator $F : X \rightarrow Y$, und Funktionalen $J : X \rightarrow \mathbb{R}$ und $G : Y \times Y \rightarrow \mathbb{R}$ welche als eigentlich, koerziv und stetig vorausgesetzt sind. Die Vergleichsanalyse der Verfahren steht nicht im Mittelpunkt, so dass die Beziehung der Ansätze zueinander nur kurz diskutiert wird.

Wenn die Abbildung $F : X \rightarrow Y$ als linear vorausgesetzt wird, so sind alle drei Minimierungsprobleme konvex, und daher stimmen ihre Lösungen für eine geeignete Wahl von α bzw. ρ überein s. [37, Sec. 3.5]. Dies ist jedoch für nichtlineare Operatoren nicht mehr der Fall, wie ein Gegenbeispiel in [41] gezeigt hat.

Ohne zusätzliche Forderungen an die Funktionenräume X und Y wurde folgendes Resultat gewonnen s. [41, Thm. 2.3; Rem. 2.6], [45, Prop. 2.5].

Satz 2.6. Sei $F : X \rightarrow Y$ stetig, nicht notwendigerweise linear, $y \in Y$ beliebig aber fest. Dann gilt

- (i) existiert eine eindeutige Lösung x_0 von (I), dann ist x_0 die eindeutige Lösung von (M) mit $G(F(x_0), y) = \delta$
- (ii) existiert eine eindeutige Lösung x_0 von (M), dann ist x_0 die eindeutige Lösung von (I) mit $J(x_0) = \rho$
- (iii) existiert eine eindeutige Lösung x_0 von (T) mit $\alpha = \alpha_0$, dann ist x_0 die eindeutige Lösung von (I) für $\rho = J(x_0)$ und die eindeutige Lösung von (M) für $\delta = G(F(x_0), y)$
- (iv) Seien darüber hinaus beide Funktionale G und J konvex, dann existiert für die eindeutige Lösung x_0 von (I) oder (M) ein $\alpha_0 > 0$, so dass x_0 die eindeutige Lösung von (T) mit $\alpha = \alpha_0$ ist.

Die Ivanov-Regularisierung ist erfolgreich für die Lösung der Identifikationsprobleme in der Naturwissenschaft eingesetzt worden [58], [65], [64], [62]. Allerdings gibt es kaum Untersuchungen, welche sich mit der optimalen Wahl des Parameters ρ beschäftigen. Ebenfalls gibt es kaum Untersuchungen zu den erschwerlichen Bedingungen in der Modellierung, wie zum Beispiel der Anwendung der Methoden auf nicht-reflexiven oder nicht-separablen Lösungsräumen.

2.3 SCHWACHE KOMPAKTHEIT IN BANACHRÄUMEN

Die Diskussion in diesem Abschnitt dient der kurzen Erläuterung von Begriffen *schwache* - und *schwach** Topologie auf Banachräumen und der grundlegenden Sätze aus der Funktionalanalysis, welche für die Zielsetzung dieser Arbeit von großer Bedeutung sind. Zunächst einige grundlegende Definitionen [1], [66].

Die Menge der Teilmengen \mathcal{T} einer Menge X mit folgenden Eigenschaften

- (i) $\emptyset \in \mathcal{T}, X \in \mathcal{T}$
- (ii) $\mathcal{T}' \subset \mathcal{T} \Rightarrow \bigcup_{O \in \mathcal{T}'} O \in \mathcal{T}$
- (iii) $O_1, O_2 \in \mathcal{T} \Rightarrow O_1 \cap O_2 \in \mathcal{T}$

heißt *Topologie* auf X , das Paar (X, \mathcal{T}) heißt *topologischer Raum* und die Elemente von \mathcal{T} nennt man *offene Mengen*. Eine Menge aus \mathcal{T} wird als *offene Umgebung* des Elements $x \in X$ bezeichnet, wenn sie x enthält. Ein topologischer Raum (X, \mathcal{T}) heißt *Hausdorffraum*, wenn für alle $x_1, x_2 \in X$ mit $x_1 \neq x_2$, offene Umgebungen O_1, O_2 existieren, so dass $O_1 \cap O_2 = \emptyset$. Ein Element $x \in X$ heißt *Häufungspunkt* der Menge $M \subset X$, wenn jede offene Umgebung von x mindestens ein Element $m \in M$ enthält mit $m \neq x$. Die Menge M heißt *abgeschlossen* bzgl. der Topologie \mathcal{T} , wenn sie alle ihre Häufungspunkte enthält. Es ist leicht zu zeigen, dass M genau dann abgeschlossen ist, wenn sie als Komplement einer offenen Menge $X \setminus O$ darstellbar ist. Eine Folge $\{x_n\} \subset X$ *konvergiert* zu $x \in X$ bzgl. der Topologie \mathcal{T} , wenn für alle offenen Umgebungen $O \in \mathcal{T}$ von x ein $N \in \mathbb{N}$ existiert, so dass für alle $n > N$ gilt $x_n \in O$. In einem Hausdorffraum ist der Grenzwert einer konvergenten Folge eindeutig bestimmt.

Jeder metrische Raum ist ein Hausdorffraum - denn man kann zeigen, dass die Definitionen des Innern und des Abschlusses einer Menge bzgl. einer Topologie und bzgl. einer Metrik übereinstimmen. Als vollständiger normierter Raum, ist der Banachraum ein topologischer Raum mit der Norm-Topologie, genauer ein lokalkonvexer Hausdorffraum. Wenn man die Elemente des Banachraumes als eine Menge betrachtet, kann man auf dieser Menge schwächere Topologie konstruieren, so dass der Raum induziert durch diese Topologie ebenfalls ein lokalkonvexer Hausdorffraum ist.

Wenn auf einer Menge X zwei verschiedene Topologien $\mathcal{T}_1, \mathcal{T}_2$ definiert sind, dann wird die Topologie \mathcal{T}_2 stärker (oder feiner) als \mathcal{T}_1 bezeichnet, bzw. \mathcal{T}_1 schwächer (oder gröber) als \mathcal{T}_2 , wenn $\mathcal{T}_1 \subset \mathcal{T}_2$. Die Motivation zur Abschwächung der Topologie liegt in der Erwartung, dass ein Raum mit der gröberen Topologie mehr kompakte Teilmengen zulässt, so dass z.B. die Einheitskugel $B_1(0)$ in (X, \mathcal{T}) kompakt ist. Ein formelles Mittel zur Konstruktion solcher Topologie liefert die Eigenschaft der stetigen Funktionalen, nämlich dass das Urbild einer offenen Menge eines stetigen Funktionalen offen ist.

So kann auf einem Banachraum E die *schwache Topologie* $\sigma(E, E')$, als die schwächste Topologie definiert werden, in der alle linearen norm-stetigen Funktionalen aus E' stetig bleiben. Analog kann auf einem Dualraum E' eines Banachraumes E die *schwach** Topologie $\sigma(E', E)$, als die schwächste Topologie definiert werden, in der alle linearen norm-stetigen Funktionalen aus E stetig bleiben [9, Kap.: 3.2; 3.4].

Zum anschaulichen Vergleich der Topologien, werden hier beide auf dem selben Banachraum E' betrachtet. E' wird gezielt gewählt, um zu betonen, dass die *schwach** Topologie

ausschließlich auf Dualräumen definiert werden kann. Dann wird durch $\sigma(E', E'')$ die schwache Topologie und durch $\sigma(E', E)$ die schwach* Topologie auf E' definiert. In reflexiven Räumen, wenn die stetige lineare Isometrie $J_E : E \rightarrow E''$ zusätzlich isomorph ist, stimmen diese beiden Topologien überein. Sonst, falls existiert, ist die schwach* Topologie größer als die schwache. Für diese, auf einem Banachraum definierten schwachen und schwach* Topologien, gelten folgende Eigenschaften [9, Prop.: 3.3; 3.5; 3.11; 3.13].

- (i) Die topologischen Räume $(E', \sigma(E', E''))$ und $(E', \sigma(E', E))$ sind lokalkonvexe Hausdorff-Räume.
- (ii) Eine Folge x'_n aus E' konvergiert gegen x' bzgl. der Topologie $\sigma(E', E'')$ genau dann, wenn gilt

$$\lim_{n \rightarrow \infty} x''(x'_n) = x''(x') \quad \text{für alle } x'' \in E''; \quad \text{bezeichnet durch } x'_n \rightarrow x'$$

- (iii) Eine Folge x'_n aus E' konvergiert gegen x' bzgl. der Topologie $\sigma(E', E)$ genau dann, wenn

$$\lim_{n \rightarrow \infty} x'(x_n) = x'(x) \quad \text{für alle } x \in E; \quad \text{bezeichnet durch } x'_n \rightharpoonup^* x'$$

Bei Kompaktheitseigenschaften im topologischen Raum (X, \mathcal{T}) wird zwischen *überdeckungskompakten* und *folgenkompakten* Mengen unterschieden. Als *kompakte* bzw. *überdeckungskompakte* Teilmenge eines topologischen Raumes wird eine Menge bezeichnet, deren offene Überdeckungen eine endliche Teilüberdeckung enthalten. Als *folgenkompakt* wird eine Teilmenge eines topologischen Raumes bezeichnet, wenn jede unendliche Folge dieser Menge eine konvergente Teilfolge besitzt. In metrischen Räumen ist eine Teilmenge kompakt genau dann, wenn sie folgenkompakt ist (Beweis s. z.B. [1]). Im Allgemeinen gilt diese Behauptung jedoch nicht, da nicht jeder topologischer Raum metrisierbar ist.

Der *Satz von Kakutani* stellt eine Verbindung der schwachen Kompaktheit der Einheitskugel und den Eigenschaften eines Banachraumes s. [9, Thm. 3.17] dar.

Satz 2.7. *Ein Banachraum E ist genau dann reflexiv, wenn die abgeschlossene Einheitskugel $B_1(0) \subset E$ überdeckungskompakt bzgl. der schwachen Topologie auf E ist.*

Eine analoge Aussage gilt mit dem *Satz von Banach–Alaoglu* für die schwach* Kompaktheit s. [1, Kap. 6.7].

Satz 2.8. *Sei E' ein Banachraum. Dann ist die abgeschlossene Einheitskugel $B_1(0) \subset E'$, überdeckungskompakt bzgl. der schwach* Topologie auf E' .*

Für diese Arbeit sind die Begriffe der schwach - und schwach* *Folgenkompaktheit* von größerer Bedeutung. Der *Satz von Eberlein–Šmulian* besagt, dass in einem Banachraum die Begriffe der schwachen Überdeckungs - und Folgenkompaktheit übereinstimmen [61, Thm. VIII.6.1].

Satz 2.9. *Sei E ein Banachraum. Die Menge $M \subset E$ ist genau dann schwach kompakt, wenn M schwach folgenkompakt ist.*

2 Grundlagen und Zielsetzung

In einer noch größeren Topologie gilt die Aussage des oberen Satzes nicht mehr. Im Falle eines separablen prädualen Raumes kann der *Satz von Banach–Alaoglu* s. [1, Satz 6.5] jedoch wie folgt formuliert werden.

Satz 2.10. *Sei E separabel. Dann ist die abgeschlossene Einheitskugel $B_1(0) \subset E'$, folgenkompakt bzgl. der schwach* Topologie auf E' .*

Diese Aussage gilt dann auch für jede andere abgeschlossene Kugel $B_r(x) \subset E'$ mit $x \in E'$ und Radius $r \geq 0$.

Zum Schluss dieses Abschnittes noch einige wichtige Folgerungen:

- In einem Banachraum besitzt jede beschränkte unendliche Folge eine schwach-konvergente Teilfolge. Jede schwach-konvergente Folge ist beschränkt.
- In einem Banachraum mit separablem Prädualem besitzt jede beschränkte unendliche Folge eine schwach*-konvergente Teilfolge. Jede schwach*-konvergente Folge ist beschränkt.

2.4 ZIELSETZUNG

Im Vordergrund der Motivation dieser Arbeit steht nun die Idee, welche im vorherigen Paragraphen bereits angesprochen wurde, und zwar eine Abschwächung der Topologie, in der Erwartung, dass eine gröbere Topologie weitere kompakte Teilmengen zulässt. Nach Satz 2.8 wird die Ivanovs Forderung nach der Kompaktheit des Domains $D(F) \subset U$ für einen Dualraum U eines Banachraumes erreicht, wenn man als Lösungsmenge die abgeschlossene Einheitskugel $B_1(0)$ betrachtet. Somit wird die Frage, ob sich die Ivanovs Regularisierungsideologie auf der schwächsten Topologie auf einem Banachraum anwenden lässt, zur Motivation dieser Arbeit.

Ivanovs Idee der Ausdehnung des kompakten Definitionsbereiches $D(F) \subset U$ als Regularisierungsstrategie wirft eine weitere Frage auf, ob einzig durch die Variation des Radius ρ der Kugel $B_\rho(0)$ eine Regularisierungsstrategie vorliegen würde. Der Urbildraum U könnte als ein Raum mit separablem Prädualraum vorausgesetzt werden. Solch eine Voraussetzung würde es erlauben, die Kompaktheit in der schwach* Topologie als die schwach* Folgenkompaktheit s. Satz 2.10 zu behandeln, um die Arbeit mit topologischen Begrifflichkeiten nicht zu erschweren.

Der Banachraum L^∞ ist für die beschriebene Zielsetzung der ideale Kandidat, da unter anderem solch eine Wahl auch Seitens praktischer Anwendbarkeit bei Parameteridentifikationproblemen von großem Interesse wäre.

Die Lösungsmenge $M_\rho \subset U \equiv L^\infty$ mit

$$(2.4.1) \quad M_\rho := \{u \in U : \|u\|_U \leq \rho\},$$

ist für jedes $\rho \geq 0$ eine schwach*-folgenkompakte Menge in U .

Die nicht-Separabilität von $L^\infty(\Omega)$ und die Separabilität seines Prädualraumes sind nicht schwer zu zeigen. Nach dem Riesz–Darstellungssatz für den Banachraum $L^\infty(\Omega, \mu)$ mit σ -endlichem Maß μ , existiert für jedes stetige lineare Funktional $\varphi \in L^1(\Omega, \mu)^*$ ein eindeutiges Element $u \in L^\infty(\Omega, \mu)$ s. [9, Thm.: 4.13, 4.14], so dass

$$\langle \varphi, g \rangle = \int ug \quad \forall g \in L^1$$

und es gilt die Isometrie

$$\|u\|_{L^\infty} = \|\varphi\|_{L^1^*}$$

Damit hat der Raum $L^\infty(\Omega) = L^1(\Omega)^*$ einen separablen Prädualraum, obwohl er selbst nicht separabel ist, was mit folgendem Beispiel gezeigt werden kann.

Sei $f_t(x)$ eine Charakteristische Funktion in $U \equiv L^\infty(\Omega)$ mit $\Omega := [0, 1]$ so, dass

$$f_t(x) := \begin{cases} 1 & \text{für } x \leq t \\ 0 & \text{für } x > t \end{cases}$$

Bei der Annahme, dass U separabel ist, muss eine abzählbare Teilmenge $Q \subset U$ existieren, so dass

$$\forall \varepsilon > 0 \quad \forall t \in [0, 1] \quad \exists q \in Q \text{ mit } \|f_t - q\|_{L^\infty} < \varepsilon$$

2 Grundlagen und Zielsetzung

Für den Abstand beliebiger Elemente $f_t \in L^\infty$ mit $t_1 \neq t_2$ gilt mit *esssup*-Norm

$$\|f_{t_1} - f_{t_2}\|_{L^\infty} = 1$$

Das heißt, dass für ein vorgegebenes $\varepsilon := 1/3$ die Kardinalität Q mindestens Kontinuum sein muss, damit es in jeder Umgebung $B_\varepsilon(f_t)$ mindestens ein Element $q \in Q$ gibt.

Bestimmte geometrische Eigenschaften der Norm, welche für die Optimierungstheorie wichtig sind, sind hier ebenfalls nicht gegeben. Die Norm in $L^\infty(\Omega)$ ist zwar *konvex*, aber nicht *strikt konvex*. Die zur strikten Konvexität duale Eigenschaft, die *Glattheit* des Raumes, ist hier sicherlich auch nicht vorhanden, da der Prädualraum $L^1(\Omega)$ selbst nicht strikt konvex ist.

Offensichtlich ist auch die Reflexivität nicht gegeben. Der Dualraum von $L^\infty(\Omega, \mu)$ ist der Raum aller beschränkter, endlich additiver Maße auf (Ω, μ) , die absolut stetig bezüglich μ sind, ausgestattet mit der totalen Variation als Norm [67, Thm. 2.3] In der Literatur bezeichnet als

$$ba(\Omega, \mu) := L^\infty(\Omega, \mu)^*$$

Es gilt folgende Inklusion

$$L^1(\Omega, \mu) \subsetneq ba(\Omega, \mu)$$

In folgenden Kapiteln sollen nun, bei den festgelegten Voraussetzungen an den Lösungsraum, grundsätzliche Fragen zum Problem (I) in Anlehnung an die klassische Regularisierungstheorie beantwortet werden:

- Welche Bedingungen garantieren die Existenz der Quasilösungen zu beliebigen Daten $y^\delta \in Y$, für einen beliebig gewählten realen Parameter $\rho > 0$?
- Welche Art der Stabilität kann gewährleistet werden ?
- Welche Regeln zu der bestmöglichen Auswahl des Regularisierungsparameters können bei der Analyse der Parameterwahlstrategien vorgeschlagen werden ?
- Welche Bedingungen garantieren die Konvergenz der Quasilösungen zu der exakten Lösung für $\delta \rightarrow 0$, und welche Konvergenzraten sind dabei erreichbar ?

Darüber hinaus werden exemplarische Beispiele jeweils für das lineare - und nicht-lineare Inverse Problem numerisch gelöst, ein optimaler Wert des Regularisierungsparameters für verschiedene Level der Störung ermittelt, um die praktische Brauchbarkeit der untersuchten Methode zu untermauern.

3 QUASILÖSUNGEN BEI LINEAREN IDENTIFIKATIONSPROBLEMEN

3.1 EXISTENZ DER QUASILÖSUNGEN

Es wird ein direktes Problem betrachtet, gegeben durch die Operatorgleichung

$$(3.1.1) \quad Au = y$$

mit einem stetigen, linearen Operator $A : U \rightarrow Y$ und Funktionenräumen U, Y vorausgesetzt wie folgt

- Y ein gleichmäßig konvexer Banachraum
- $U \equiv L^\infty(\Omega)$, und $\Omega \subset \mathbb{R}^N$ ein beschränktes Gebiet

Es wird angenommen, dass die exakten Daten $y \in R(A) \subset Y$ im Bild des Operators liegen. Vorhanden sind jedoch ausschließlich durch ein Rauschen gestörte Daten $y^\delta \in Y$. Weiterhin wird angenommen, dass der Störungslevel $\delta \geq 0$ angegeben werden kann, so dass gilt

$$\|y^\delta - y\|_Y \leq \delta$$

Das betrachtete Funktional $J : U \rightarrow \mathbb{R} \cup \{\infty\}$ in der verallgemeinerten Form der Ivanov-Regularisierung (I) wird bei der Definition der Lösungsmenge M_ρ nun in Betracht der Aufgabenstellung als *esssup*-Norm gesetzt

$$(3.1.2) \quad J(u) := \|u\|_{L^\infty}$$

Das Optimierungsproblem (I) erhält damit folgende Form

$$(3.1.3) \quad \inf_{u \in M_\rho} \|Au - y\|_Y \quad M_\rho := \{u \in L^\infty(\Omega) : \|u\|_{L^\infty} \leq \rho\}$$

Annahme 3.1. Sei die stetige, lineare Abbildung $A : U \rightarrow Y$ zusätzlich

- injektiv,*
- schwach*-abgeschlossen, d.h. für jede Folge $\{u_n\}$ mit $u_n \rightharpoonup^* u$ in U und $Au_n \rightarrow y$ in Y folgt $y = Au$.*
- Sei weiterhin das Bild des Operators $R(A)$ dicht in Y .*

3 Quasilösungen bei linearen Identifikationsproblemen

Diese Voraussetzungen zu der Abbildung sollen über das gesamte Kapitel beibehalten werden.

Die Bildmenge

$$(3.1.4) \quad A(M_\rho) = \{Au \in Y : u \in M_\rho\} =: Q_\rho$$

erhält somit wichtige Eigenschaften, welche im nächsten Lemma diskutiert werden.

Lemma 3.2. *Für jedes $\rho \geq 0$ ist die Bildmenge Q_ρ nicht leer, beschränkt, konvex and abgeschlossen.*

Beweis. $A : U \rightarrow Y$ ist linear und somit ist $0 \in Q_\rho$ für jedes $\rho \geq 0$.

$A : U \rightarrow Y$ ist als linearer und stetiger Operator beschränkt. Es existiert also ein $C > 0$, so dass für alle $Au \in Q_\rho$ gilt

$$\|Au\|_Y \leq \|A\|_{\mathcal{L}(U,Y)} \|u\|_U \leq C\rho$$

Die abgeschlossene Kugel $M_\rho \subset L^\infty(\Omega)$ ist konvex, so dass $u := tu_1 + (1-t)u_2 \in M_\rho$ $t \in [0, 1]$ für jedes $u_1, u_2 \in M_\rho$. Weiterhin liefert die Linearität der Abbildung

$$tAu_1 + (1-t)Au_2 = A(tu_1 + (1-t)u_2) = Au \in Q_\rho \quad q_1, q_2 \in Q_\rho, \quad t \in [0, 1].$$

Sei $\{y_n\}_{n \in \mathbb{N}}$ eine Folge in $Q_\rho \subset Y$, welche stark zu einem beliebigen $y \in Y$ konvergiert. Nach der Definition von Q_ρ existiert für jedes $n \in \mathbb{N}$ ein $u_n \in M_\rho$, so dass $Au_n = y_n$. Da $\{u_n\}_{n \in \mathbb{N}} \subset M_\rho$ beschränkt ist und U ein Dualraum eines separablen Vektorraumes ist, kann das Banach–Alaoglu Theorem angewendet werden. Das heißt, es existiert eine Teilfolge $\{u_k\} \subset \{u_n\}$, so dass $u_k \rightharpoonup^* u \in M_\rho$. Mit der Forderung an die Abbildung $A : U \rightarrow Y$ sei schwach* – abgeschlossen folgt $y = Au \in Q_\rho$ und somit die Abgeschlossenheit der Menge Q_ρ für jedes $\rho \geq 0$. \square

Folgerung 3.3. *Mit den Voraussetzungen an die Abbildung Annahme 3.1 ist die Bildmenge Q_ρ schwach abgeschlossen.*

Beweis. Sei $y \in Y \setminus Q_\rho$ beliebig. Da die Menge Q_ρ nicht leer, beschränkt, konvex and abgeschlossen ist, gibt es nach dem Trennungssatz von Hahn–Banach für lokalkonvexe Räumen ein $y^* \in Y^*$ und ein $\lambda \in \mathbb{R}$, so dass

$$\langle y^*, y \rangle < \lambda < \langle y^*, q \rangle \quad \forall q \in Q_\rho$$

Die Menge Q_ρ enthält also alle schwachen Häufungspunkte s., z.B., [6]. \square

Satz 3.4. *Für jedes $y \in Y$ und jedes $\rho \geq 0$ existiert eine eindeutige Lösung $u_\rho \in M_\rho$ zu (3.1.3).*

Beweis. Zum Nachweis wird anstatt der Urbildmenge die Bildmenge Q_ρ in Y betrachtet. Das Funktional $\|Au - y\|_Y$ ist als Norm in Y schwach unterhalbstetig. Mit Annahme 3.1 ist Y reflexiv (nach dem Satz von Milman–Pettis), die Menge Q_ρ beschränkt und abgeschlossen. Es existiert also ein Minimierer $\tilde{q} \in Q_\rho$ nach der direkten Methode der Variationsrechnung

s., z.B., [21], [24]. Für $\tilde{q} \in Q_\rho$ existiert wegen der Injektivität des Operators ein eindeutiges Urbildelement $\tilde{u} \in M_\rho$ mit $\tilde{q} = A\tilde{u}$, so dass

$$\|A\tilde{u} - y\| = \inf \{ \|Au - y\| : u \in M_\rho \}$$

Seien $u_1, u_2 \in M_\rho$ mit $u_1 \neq u_2$ Lösungen zum Problem (3.1.3). Mit der injektiven Abbildung gilt für $q_1 := Au_1$ und $Au_2 := q_2$ ebenfalls $q_1 \neq q_2$, so dass

$$d = \|q_1 - y\| = \|q_2 - y\| = \inf \{ \|q - y\| : q \in Q_\rho \}$$

Für Elemente in Y gilt mit der Annahme 3.1 folgende Eigenschaft

$$\text{falls gilt } \|v_1\| = \|v_2\| = \left\| \frac{v_1 + v_2}{2} \right\| \text{ dann gilt } v_1 = v_2$$

Seien $v_1 := q_1 - y$, $v_2 := q_2 - y$. Mit $\frac{q_1 + q_2}{2} \in Q_\rho$ folgt

$$d \leq \left\| \frac{q_1 + q_2}{2} - y \right\| = \frac{1}{2} \|(q_1 - y) + (q_2 - y)\| \leq \left\| \frac{y - q_1}{2} \right\| + \left\| \frac{y - q_2}{2} \right\| = d$$

Damit ein Widerspruch zu der Annahme $q_1 \neq q_2$ und die Eindeutigkeit der Lösung. □

3.2 REGULARISIERUNG

Der Nachweis der Existenz von Lösungen zum Problem (3.1.3) macht es möglich eine Näherungsweise-Lösungsmethode vorschlagen zu können. Damit aber die Lösungen $\{u(\rho, y)\}_{\rho>0}$ mit ρ - als Regularisierungsparameter ein Regularisierungsverfahren darstellen, muss eine Reihe von Grenzwertbeziehungen erfüllt sein. In dieser Arbeit dienen die Definitionen einer Regularisierung aus [29], [20] als Richtlinien. Das klassische Regularisierungsschema angewandt an das vorliegende Optimierungsproblem, muss dabei aber an die Voraussetzungen (4.1) angepasst werden, denn eine starke Konvergenz der Lösungen kann hier nicht erwartet werden.

Definition 3.5. Die Quasilösungen $\{u(\rho, y)\}_{\rho>0}$ mit ρ - als Regularisierungsparameter bilden eine konvergente Regularisierung, wenn folgende drei Bedingungen erfüllt sind:

- i) $u(\rho, y_n) \xrightarrow{n \rightarrow \infty}^* u(\rho, y)$ für jedes $\rho \geq 0$ und $y_n \xrightarrow{n \rightarrow \infty} y$
- ii) $u(\rho_n, y) \xrightarrow{n \rightarrow \infty}^* A^{-1}y$ für jedes $y \in R(A)$ und $\rho_n \xrightarrow{n \rightarrow \infty} \infty$
- iii) $\exists (\delta, y^\delta) \mapsto \rho(\delta, y^\delta) : u(\rho(\delta, y^\delta), y^\delta) \xrightarrow{*} u^\dagger$ für $\delta \rightarrow 0$

dabei definiert die Funktion $\rho : \mathbb{R}^+ \times Y \rightarrow \mathbb{R}^+$; $(\delta, y^\delta) \mapsto \rho(\delta, y^\delta)$ eine Parameterwahlstrategie.

Mitunter der Existenz und Eindeutigkeit der Lösung zeigt Satz 3.4: - für jedes $\rho > 0$ und jedes $y \in Y$ existiert eine eindeutige metrische Projektion $y \mapsto P_{Q_\rho}(y) =: q_\rho$

$$(3.2.1) \quad \|q_\rho - y\|_Y = \min_{q \in Q_\rho} \|q - y\|_Y$$

Die Stetigkeit des Projektionsoperators $P : Y \rightarrow Q_\rho$ ist wesentlich um die *Stabilität*, also den ersten Punkt der Definition 3.5 nachzuweisen.

Proposition 3.6. Sei Q_ρ nicht leer, konvex und abgeschlossen und Y ein gleichmäßig konvexer Banachraum. Dann ist der Projektionsoperator $P : Y \rightarrow Q_\rho$ auf dem ganzen Raum Y definiert, eindeutig und stetig. Ist Y darüber hinaus ein Hilbertraum, so ist die Projektion Lipschitzstetig mit der Lipschitzkonstante 1.

Beweis. Die Existenz und Eindeutigkeit der Projektion folgt direkt aus Satz 3.4. Weiterhin gilt für $y_n \rightarrow y$ (s., z.B., [37, sec. 1.6.]

$$\begin{aligned} \lim_{n \rightarrow \infty} \|P(y_n) - y\| &\leq \lim_{n \rightarrow \infty} \|P(y_n) - y_n\| + \|y_n - y\| \\ &\leq \lim_{n \rightarrow \infty} \|P(y) - y_n\| + \|y_n - y\| \\ &= \|P(y) - y\| \end{aligned}$$

Nach dem Satz von Milman–Pettis ist Y reflexiv und $Q_\rho \in Y$ beschränkt und schwach abgeschlossen (Folgerung 3.3). Somit besitzt die Folge $\{q_n := P(y_n)\}$ nach Satz von Eberlein–Šmulian mindestens einen schwachen Häufungspunkt $q \in Q_\rho$ und eine Teilfolge $q_{n_m} \rightharpoonup q$.

Die gleichmäßig konvexen Räume erfüllen die Radon-Riesz-Eigenschaft, so dass für

$$\|q_{n_m} - y\| \rightarrow \|q - y\|$$

die starke Konvergenz $q_{n_m} \rightarrow q$ gilt. Da diese Behauptung für alle möglichen schwachen Häufungspunkte der Folge $\{q_n := P(y_n)\}$ gilt und die Projektion eindeutig ist, so gilt für die gesamte Folge

$$P(y_n) \rightarrow q = P(y)$$

In einem Hilbertraum ist für die metrische Projektion auf eine konvexe Teilmenge folgende Ungleichung gültig s., z.B., [5]

$$(3.2.2) \quad \langle q - P(y), y - P(y) \rangle \leq 0 \quad \forall q \in Q_\rho$$

Sei $q \in Q_\rho$ aus (3.2.2) wie folgt gesetzt: $q_1 := P(y_1), q_2 := P(y_2)$. So ergibt sich ein Ungleichungssystem

$$\begin{aligned} \langle q_1 - P(y_2), y_2 - P(y_2) \rangle &\leq 0 \\ \langle q_2 - P(y_1), y_1 - P(y_1) \rangle &\leq 0 \end{aligned}$$

Das Addieren und Umformen ergibt

$$\begin{aligned} &\langle P(y_1) - P(y_2), y_2 - P(y_2) \rangle + \langle P(y_2) - P(y_1), y_1 - P(y_1) \rangle \leq 0 \\ \Rightarrow &\|P(y_1) - P(y_2)\|^2 - \langle P(y_1) - P(y_2), y_1 - y_2 \rangle \leq 0 \\ \Rightarrow &\|P(y_1) - P(y_2)\|^2 \leq \|P(y_1) - P(y_2)\| \|y_1 - y_2\| \\ \Rightarrow &\|P(y_1) - P(y_2)\| \leq \|y_1 - y_2\| \end{aligned}$$

□

Nun kann die stabile Abhängigkeit der Lösungen von den Eingangsdaten im Sinne der Definition 3.5(i) gezeigt werden.

Satz 3.7. Seien $\rho \geq 0$ und $y \in Y$ gegeben. Sei $\{y_n\}_{n \in \mathbb{N}} \subset Y$ eine beliebige Folge mit $y_n \rightarrow y$. Dann konvergiert die entsprechende Folge von Lösungen $\{u_n := u(\rho, y_n)\}_{n \in \mathbb{N}}$ schwach* zu

$$u_\rho := u(\rho, y) = \arg \min_{u \in M_\rho} \|Au - y\|$$

Beweis. Nach Satz 3.4 existiert zu jedem $n \in \mathbb{N}$ eine eindeutige metrische Projektion $q_n \in Q_\rho$. Weiterhin folgt mit Proposition 3.6 $q_n \rightarrow q := P_{Q_\rho} y$.

Wie im Nachweis zur Abgeschlossenheit von Q_ρ gezeigt s. Lemma 3.2, kann eine Teilfolge $\{u_k\}$ extrahiert werden, so dass $u_k \rightharpoonup^* u \in M_\rho$ und $Au = q$, d.h. $u \in M_\rho$ minimiert $\|Au - y\|_Y$ auf M_ρ . Das zeigt, dass $u = u_\rho$ die Lösung des Problems (3.1.3) ist. Die Eindeutigkeit dieser Lösung folgt mit Satz 3.4.

Um zu zeigen, dass die gesamte Lösungsfolge gegen denselben Grenzwert konvergiert, wird die Menge der restlichen Elemente der Folge $\{u_n\} \setminus \{u_k\}$ betrachtet, welche ebenfalls beschränkt in M_ρ ist. Das heißt, dass diese Menge ebenfalls mindestens einen schwach* Häufungspunkt besitzt. Die Annahme, dass solch ein Häufungspunkt von u_ρ verschieden ist, führt mit der Forderung an die Abbildung (schwach*-abgeschlossen) und der starken Konvergenz der Bildfolge $Au_n \rightarrow q = Au_\rho$ sofort zum Widerspruch. Somit kann es keinen von u_ρ verschiedenen Häufungspunkt geben. □

3 Quasilösungen bei linearen Identifikationsproblemen

Wie kann eine Regularisierung s. z.B., [29, Def. 4.1.] für das vorliegende Problem (3.1.3) aufgebaut werden? In [20, Prop. 3.4.], wird von einer Familie regularisierender Operatoren für $\alpha \rightarrow 0$ die punktweise Konvergenz gefordert. Analog zu dieser klassischen Vorgehensweise sollen hier die Approximationseigenschaften der Quasilösungsmethode für jedes beliebige $y \in R(A)$ gezeigt werden.

Es ist eigentlich offensichtlich, dass die Folge der Quasilösungen bei $\rho \rightarrow \infty$ für jedes $y \in R(A)$ sogar stark konvergiert. Da diese Eigenschaft in 3.5(ii) definiert ist, wird sie der Vollständigkeit halber in einem Satz behandelt.

Satz 3.8. *Sei $y \in R(A)$ gegeben, so dass $A\hat{u} = y$. Dann konvergieren für jede Folge $\{\rho_n\}_{n \in \mathbb{N}}$ mit $\rho_n \rightarrow \infty$ die entsprechenden Lösungen u_{ρ_n} stark gegen \hat{u} .*

Beweis. Die Menge Q_ρ ist konvex und hat der Definition nach die Eigenschaft $Q_{\rho_1} \subset Q_{\rho_2}$ für $\rho_1 < \rho_2$. Für eine monoton wachsende Parameterfolge $\{\rho_n\}$ existiert demnach ein $N \in \mathbb{N}$ und ein ρ_N , so dass gilt $y \in Q_{\rho_N}$ und somit für alle $n \geq N$

$$(3.2.3) \quad \inf_{u \in M_{\rho_n}} \|Au - y\| = 0 = \|A\hat{u} - y\|$$

Damit ist die Menge der entsprechenden Lösungen $\{u_1, \dots, u_N\}$ endlich, so dass

$$\lim_{n \rightarrow \infty} u_n = \hat{u}$$

□

3.3 DISTANZFUNKTION

Die von Ivanov vorgeschlagene Regularisierungsidee beruht auf der Einschränkung des Lösungsraumes zwecks Erhalt einer kompakten Lösungsmenge im Definitionsbereich des Operators. Typischerweise wird eine Kugel mit festem Radius betrachtet. Die Genauigkeit der Lösung hängt direkt vom gewählten Radius der Kugel ab, womit er zum Regularisierungsparameter wird. Auf diese Weise wird das Verhältnis zwischen Eingangsdaten und regularisierten Lösungen über die Bilder dieser Lösungen bzw. über deren Distanz zu den vorhandenen Daten beschrieben. Das heißt, dass bei der Bestimmung einer Quasilösung zu den Eingangsdaten das Verhalten der Distanzfunktion im Vordergrund steht und untersucht werden muss, damit Vorschläge zur Parameterwahl gemacht werden können. Dazu soll als Erstes gezeigt werden, dass zu jedem Element $y \in Y$ ein Parameter ρ existiert, so dass die Beziehung der Eingangsdaten $y \notin R(A)$ und der Lösungen von (3.1.3) über die *Distanzfunktion*

$$(3.3.1) \quad d : [0, \infty) \times Y \rightarrow [0, \infty) \quad (\rho, y) \mapsto \|Au_\rho - y\|_Y = \inf_{u \in M_\rho} \|Au - y\|_Y$$

beschrieben werden kann.

Die Aussage des nächsten Satzes ist zum Aufbau eines Regularisierungsverfahrens nach der Ivanovschen Ideologie wesentlich. Diese Aussage wurde für reflexive Banachräume für eine kompakte und konvexe Lösungsmenge nachgewiesen [19]. Trotz schwächerer Forderungen an die Lösungsmenge M_ρ und Funktionenraum U wird hier die selbe Beweisidee verfolgt.

Satz 3.9. *Sei Y ein reflexiver Banachraum. Für jedes $y \in Y$ ist die Abbildung $\rho \mapsto d(\rho, y)$ surjektiv auf dem offenen Intervall $(0, \|y\|_Y)$.*

Beweis. Seien $y \in Y$ und $\sigma \in (0, \|y\|_Y)$ gegeben. Es wird gezeigt, dass dann ein eindeutiges $\rho(\sigma, y)$ existiert mit

$$\|q(\sigma, y) - y\| = \min_{q \in Q_\rho} \|q - y\| = \sigma$$

Dazu soll im ersten Schritt des Beweises folgende Konstruktion durchgeführt werden:

Das Bild des Operators $R(A) \subset Y$ ist dicht (nach Annahme 3.1), so existiert ein hinreichend großes ρ , so dass $Q_\rho \cap B_\sigma(y)$ nicht leer ist. Daher existiert für alle solche $\rho > 0$, welche einen nicht leeren Durchschnitt bilden ein Infimum

$$(3.3.2) \quad \tilde{\rho} := \inf\{\rho > 0 : Q_\rho \cap B_\sigma(y) \neq \emptyset\}.$$

Nach Definition existiert also eine Folge $\{\rho_n\}_{n \in \mathbb{N}} \subset (0, \infty)$, so dass

$$S_n := Q_{\rho_n} \cap B_\sigma(y) \neq \emptyset$$

und $\rho_n \rightarrow \tilde{\rho}$. O.B.d.A. kann angenommen werden, dass diese Folge monoton fallend ist, so dass $Q_{\rho_{n+1}} \subseteq Q_{\rho_n}$ und somit $S_{n+1} \subseteq S_n$.

Im zweiten Schritt des Beweises wird gezeigt, dass die Schnittmenge $S := \bigcap_{n \in \mathbb{N}} S_n$ nicht leer ist, und für alle Elemente $s \in S$ auch $s \in Q_{\tilde{\rho}}$ erfüllt wird:

Für jedes $n \in \mathbb{N}$ ist $S_n \subset Y$ beschränkt, konvex und abgeschlossen als Durchschnittsmenge solcher und nach der Konstruktion nicht leer. So können für jedes $n \in \mathbb{N}$ Projektionen

3 Quasilösungen bei linearen Identifikationsproblemen

$s_n := P_{S_n}(0)$ betrachtet werden. Die Monotonie von $\{\rho_n\}_{n \in \mathbb{N}}$ impliziert $\|s_n\|_Y \leq \rho_1$. Daher ist $\{s_n\}_{n \in \mathbb{N}}$ beschränkt und besitzt in einem reflexiven Banachraum eine schwach konvergente Teilfolge $s_k \rightharpoonup s$ mit einem schwachen Grenzwert $s \in Y$. Der Durchschnitt beliebig vieler abgeschlossener konvexer Mengen $\bigcap_{n \in \mathbb{N}} S_n$ ist selbst abgeschlossen und konvex. Weiterhin gilt, dass abgeschlossene konvexe Mengen schwach abgeschlossen sind, so dass $s \in \bigcap_{k \in \mathbb{N}} S_k$ für jedes $k \in \mathbb{N}$, das heißt $s \in S$.

Nach der Konstruktion gilt $s \in S_n \subset Q_{\rho_n}$ für alle $n \in \mathbb{N}$ und somit das Urbildelement $A^{-1}s \in M_{\rho_n}$ für alle $n \in \mathbb{N}$ (unter anderem $\|A^{-1}s\| \leq \rho_n$ nach Definition), so dass bei $\rho_n \rightarrow \tilde{\rho}$ folgt $\|A^{-1}s\| \leq \tilde{\rho}$ und somit $A^{-1}s \in M_{\tilde{\rho}}$. Im Grunde gilt sogar

$$(3.3.3) \quad \|A^{-1}s\|_U = \tilde{\rho}$$

denn bei der Annahme die Ungleichung wäre streng, existiert ein $\hat{\rho} < \tilde{\rho}$, so dass $s \in Q_{\hat{\rho}} \cap B_\sigma(y)$ im Widerspruch zu der Definition von $\tilde{\rho}$ (3.3.2).

Im dritten Schritt des Beweises wird gezeigt, dass $\sigma = \|s - y\|$:

Nach der Konstruktion der Schnittmenge ist $s \in B_\sigma(y)$, denn $s \in S_n$ für alle $n \in \mathbb{N}$. Sei angenommen, dass $\|s - y\|_Y < \sigma$. Diese Annahme bedeutet, dass s ein innerer Punkt von $B_\sigma(y)$ ist. Das heißt, dass für hinreichend kleines $\varepsilon > 0$ ein $\hat{s} \in S$ existiert mit $\|\hat{s} - s\| < \varepsilon$. (Beachte: $0 < \sigma < \|y\|_Y$ impliziert $0 \notin B_\sigma(y)$ und $s \neq 0$.) A ist ein linearer Operator, so dass $0 \in Q_{\tilde{\rho}}$. Weiterhin ist $Q_{\tilde{\rho}}$ konvex und abgeschlossen Lemma 3.2, so dass für ein $\hat{s} := \varepsilon 0 + (1 - \varepsilon)s$ mit $\varepsilon > 0$ hinreichend klein, gilt $\hat{s} \in Q_{\tilde{\rho}}$. Damit ergibt sich folgende Ungleichung

$$\|A^{-1}\hat{s}\|_U = (1 - \varepsilon)\|A^{-1}s\|_U = (1 - \varepsilon)\tilde{\rho} < \tilde{\rho}$$

im Widerspruch zu der Definition von $\tilde{\rho}$ (3.3.2). Das heißt $\|s - y\|_Y = \sigma$ für jedes $s \in S$.

Um die Beweisführung abzuschließen sei angenommen, dass es ein $\bar{q} \in Q_{\tilde{\rho}}$ existiert mit $\|\bar{q} - y\|_Y < \sigma$. Damit gibt es wiederum ein $\hat{q} \in Q_{\tilde{\rho}}$ und $\hat{q} \in B_\sigma(y)$, so dass $\|A^{-1}\hat{q}\|_U < \tilde{\rho}$ im Widerspruch zu Definition von $\tilde{\rho}$ (3.3.2). \square

Im Weiteren sollen noch kurz einige Bemerkungen zum obigen Satz diskutiert werden.

- Im Prinzip könnte der obere Beweis aus dem *Cantor-Intersection-Theorem* gefolgert werden. Es wäre zu zeigen, dass der Durchmesser der konstruierten Mengen S_n für $n \rightarrow 0$ gegen Null strebt. Doch dazu wäre die Forderung Y sei ein gleichmäßig konvexer Banachraum nötig, welche in Satz 3.9 jedoch auf reflexive Räume abgeschwächt werden konnte.
- In der Beweisführung musste nicht explizit gezeigt werden, dass $s \in S$ ein eindeutiges Repräsentant der Schnittmenge $S := \bigcap_{n \in \mathbb{N}} S_n$ ist. Diese Eindeutigkeit folgt jedoch bei der Forderung Y sei gleichmäßig konvex aus der Konvexität der Menge Q_ρ Lemma 3.2.

Da das Verhalten der Distanzfunktion von solch großer Bedeutung für diese Arbeit ist, werden hier im Anschluss an Satz 3.9 weitere Eigenschaften diskutiert.

- $d(0, y) = \|y\|_Y$ - folgt direkt aus (3.3.1).
- Für $y \notin R(A)$ gilt $d(\rho, y) \rightarrow 0$ für $\rho \rightarrow \infty$. Denn der Operator hat nach Annahme 3.1 ein dichtes Bild $R(A) \subset Y$, so dass gilt $\|Au - y\| \neq 0$ für jedes $\rho \geq 0$.

- Für $y \in R(A)$ existiert ein $\rho > 0$, so dass $d(\rho, y) = 0$, denn für hinreichend großes ρ existiert ein $u \in M_\rho \subset D(A)$, so dass $Au = y$.

Weiterhin gilt nach der Beweisführung (3.3.3) $\|u_\rho\|_U = \rho$ für alle Lösungen $u_\rho \in M_\rho$. Eine Ausnahme bildet der Fall $d(\rho, y) = 0$, d.h., $y \in Q_\rho \subset R(A)$. Es ist eine wichtige Eigenschaft der Lösungen, welche ebenfalls in [45, Prop. 2.2] bei der Voraussetzung U und Y seien Hilberträume gezeigt wurde und wird daher explizit im nächsten Korollar festgehalten.

Korollar 3.10. *Zu jedem $\rho \geq 0$ und jedem $y \in Y \setminus Q_\rho$ besitzt eine Lösung des Problems (3.1.3) folgende Eigenschaft*

$$(3.3.4) \quad \|u_\rho\|_U = \rho$$

Eine weitere bedeutende Eigenschaft ist die *strenge Monotonie* der Distanzfunktion für jedes fixierte $y \in Y$.

Korollar 3.11. *Die Abbildung $\rho \mapsto d(\rho, y)$ ist für jedes $y \in Y \setminus R(A)$ streng monoton fallend und stetig.*

Beweis. Sei $\rho_1 < \rho_2$. Dann folgt $M_{\rho_1} \subset M_{\rho_2}$ und mit (3.3.1)

$$d(\rho_2, y) = \inf_{u \in M_{\rho_2}} \|Au - y\|_Y \leq \inf_{u \in M_{\rho_1}} \|Au - y\|_Y = d(\rho_1, y).$$

Die Annahme $d(\rho_1, y) = d(\rho_2, y)$ führt zum Widerspruch, denn mit Korollar 3.10 folgt

$$\|u_{\rho_1}\|_U = \rho_1 < \rho_2 = \|u_{\rho_2}\|_U$$

und damit $u_{\rho_1} \neq u_{\rho_2}$. Und mit der Injektivität von A auch $Au_{\rho_1} \neq Au_{\rho_2}$.

Sei $\bar{u} := \frac{1}{2}(u_{\rho_1} + u_{\rho_2}) \in M_{\rho_2}$, dann folgt in einem gleichmäßig konvexen Raum Y

$$\|A\bar{u} - y\| = \left\| \frac{1}{2}(Au_{\rho_1} + Au_{\rho_2}) - y \right\|_Y < \frac{1}{2} (\|Au_{\rho_1} - y\|_Y + \|Au_{\rho_2} - y\|_Y) = d(\rho_2, y),$$

und weil $d(\rho_1, y) = d(\rho_2, y)$ vorausgesetzt wurde $\|A\bar{u} - y\| < d(\rho_2, y)$ im Widerspruch zur (3.3.1).

Die Monotonie und Surjektivität liefern nach einem bekannten Satz der Analysis die Stetigkeit der Distanzfunktion. \square

O.B.d.A. ist im Korollar 3.11 der Fall $y \in R(A) \subset Y$ ausgeschlossen worden, da sonst die *strenge Monotonie* für alle $\rho \geq \|u\|_\rho$ nicht mehr gegeben ist, denn es gilt

$$d(\rho, y) = \min_{u \in M_\rho} \|Au - y\|_Y = 0 \quad \forall \rho \geq \|u_\rho\|$$

Satz 3.12. *Für jedes $y \in Y$ ist die Abbildung $\rho \mapsto d(\rho, y)$ konvex.*

3 Quasilösungen bei linearen Identifikationsproblemen

Beweis. Seien $0 < \rho_1 < \rho_2$ gegeben. Es ist zu zeigen, dass für $\{\rho = t\rho_1 + (1-t)\rho_2 : t \in [0, 1]\}$ gilt

$$d(\rho) = d(t\rho_1 + (1-t)\rho_2) \leq td(\rho_1) + (1-t)d(\rho_2)$$

Zur $\rho_1 > 0, \rho_2 > 0$ existieren nach Satz 3.9 eindeutige Projektionen $s_1 \in \partial Q_{\rho_1}$ und $s_2 \in \partial Q_{\rho_2}$, so dass

$$\begin{aligned} d(\rho_1, y) &:= \|s_1 - y\| = \min\{\|q - y\| : q \in Q_{\rho_1}\} \\ d(\rho_2, y) &:= \|s_2 - y\| = \min\{\|q - y\| : q \in Q_{\rho_2}\} \end{aligned}$$

Betrachte zum Nachweis die Verbindungsstrecke $ts_1 + (1-t)s_2$ für $t \in [0, 1]$.

Nach Lemma 3.2 ist die Menge Q_ρ konvex, so dass für alle $0 \leq \rho \leq \rho_2$ gilt $Q_\rho \subset Q_{\rho_2}$, womit auch $s := ts_1 + (1-t)s_2 \in Q_{\rho_2}$ für alle $t \in [0, 1]$.

Es existiert ein eindeutiges Urbild $w = A^{-1}s \in U$ mit

$$\|w\| = \|A^{-1}s\| = \|tA^{-1}s_1 - (1-t)A^{-1}s_2\| \leq t\rho_1 + (1-t)\rho_2 = \rho$$

Die Distanzfunktion ist monoton fallend, so dass gilt $d(\rho, y) \leq d(\|w\|, y)$. Weiterhin gilt

$$d(\|w\|, y) = \inf\{\|Au - y\|_Y : \|u\| \leq \|w\|\}$$

und damit

$$\begin{aligned} d(\|w\|, y) &= \|Aw - y\| \\ &= \|ts_1 - (1-t)s_2 - y\| \\ &\leq t\|s_1 - y\| + (1-t)\|s_2 - y\| \\ &= td(\rho_1) + (1-t)d(\rho_2) \end{aligned}$$

und damit die Behauptung. □

Zum Schluss dieses Abschnittes soll noch eine starke Eigenschaft der Distanzfunktion gezeigt werden und zwar in ihrem zweiten Argumenten $y \in Y$.

Satz 3.13. Für jedes $\rho \geq 0$ ist die Abbildung $y \mapsto d(\rho, y)$ Lipschitzstetig mit Konstante 1.

Beweis. Betrachte $y_1, y_2 \in Y$. Für entsprechende Projektionen $q_1 := P_{Q_\rho}y_1$ und $q_2 := P_{Q_\rho}y_2$ gilt

$$\begin{aligned} d(\rho, y_1) &= \|q_1 - y_1\|_Y \leq \|q_2 - y_1\|_Y \leq \|q_2 - y_2\|_Y + \|y_2 - y_1\|_Y \\ d(\rho, y_2) &= \|q_2 - y_2\|_Y \leq \|q_1 - y_2\|_Y \leq \|q_1 - y_1\|_Y + \|y_1 - y_2\|_Y \end{aligned}$$

$$\text{und damit} \quad |d(\rho, y_1) - d(\rho, y_2)| \leq \|y_1 - y_2\|_Y \quad \square$$

3.4 PARAMETERWAHL

Nun können die klassischen Parameterwahlstrategien s. Definition 2.2 angewandt auf die Ivanov-Regularisierung, untersucht werden. Das primäre Ziel dieses Paragraphen ist es jedoch, eine Regel für die Parameterwahl zu finden, welche im Allgemeinen eine sichere Konvergenz des vorliegenden Verfahrens liefert. Daher wird eine gründliche Untersuchung bestimmter Strategien hier verworfen. Eine ausgiebige Diskussion vieler erfolgreicher Parameterwahlstrategien ist z.B. in [4] gegeben. Folgende Voraussetzungen an die Daten und die Datenfehler werden für die vorliegende Analyse getroffen.

Es wird eine nicht-negative Folge $\{\delta_n\}_{n \in \mathbb{N}}$ betrachtet mit $\delta_n \rightarrow 0$ für $n \rightarrow \infty$, und $y^{\delta_n} \in Y$ mit $\|y^{\delta_n}\|_Y > 0 \quad \forall n \in \mathbb{N}$, so dass

$$\|y^{\delta_n} - y\|_Y \leq \delta_n \quad \text{für jedes } n \in \mathbb{N}$$

Weiterhin wird vorausgesetzt, dass die exakten Daten $Au^\dagger = y \in R(A)$ stets im Bild des Operators liegen.

Die Lösungen zu den Daten y^{δ_n} werden jeweils durch

$$(3.4.1) \quad u_{\rho_n}^{\delta_n} = \arg \min_{u \in M_{\rho_n}} \|Au - y^{\delta_n}\|_Y$$

und die eindeutige Quasilösung zu den exakten Daten in (3.1.1) durch $u^\dagger \in U$ bezeichnet.

Im nächsten Schritt wird untersucht, welche fundamentale Eigenschaften die Regularisierungsparameter zu der vorliegenden Definition 3.5 im Allgemeinen erfüllen müssen. Der nächste Satz liefert die Notwendigen und die Hinreichenden Bedingungen dazu.

Satz 3.14. *Sei die inverse Abbildung A^{-1} nicht stetig. Dann konvergiert für $n \rightarrow \infty$ die Folge der Lösungen $\{u_{\rho_n}^{\delta_n}\}_{n \in \mathbb{N}}$ schwach* gegen u^\dagger genau dann, wenn folgende Bedingungen erfüllt sind*

(i) $\{\rho_n\}_{n \in \mathbb{N}}$ ist beschränkt

(ii) $\liminf_{n \rightarrow \infty} \rho_n \geq \|u^\dagger\|_U$

Beweis. Die Bedingungen sind hinreichend, denn (i) impliziert die Beschränktheit der Lösungen $\{u_{\rho_n}^{\delta_n}\}_{n \in \mathbb{N}}$, so dass eine schwach*-konvergente Teilfolge extrahiert werden kann (mit den selben Indizes bezeichnet) $u_{\rho_n}^{\delta_n} \rightharpoonup^* u$ für ein $u \in U$. Mit (ii) kann im nächsten Schritt für entsprechende Parameter ρ_n eine weitere Teilfolge extrahiert werden, so dass gilt

$$\rho_{n_k} \rightarrow \hat{\rho} \quad \text{für ein } \hat{\rho} \geq \|u^\dagger\|_U$$

Zu Übersichtlichkeit soll auch diese Teilfolge im Weiteren mit den selben Indizes $n \in \mathbb{N}$ bezeichnet werden. Mit $Au^\dagger = y \in Q_{\hat{\rho}}$ und der Stetigkeit der Distanzfunktion Korollar 3.11 folgt

$$(3.4.2) \quad \|P_{Q_{\rho_n}} y - y\|_Y = d(\rho_n, y) \rightarrow d(\hat{\rho}, y) = 0$$

3 Quasilösungen bei linearen Identifikationsproblemen

Weiterhin liefert die Existenz eindeutiger metrischer Projektion $P_{Q_{\rho_n}} y^{\delta_n} = Au_{\rho_n}^{\delta_n}$ zu jedem solchen n folgende Ungleichung

$$\|P_{Q_{\rho_n}} y^{\delta_n} - y^{\delta_n}\|_Y = \inf_{q \in Q_{\rho_n}} \|q - y^{\delta_n}\|_Y \leq \|P_{Q_{\rho_n}} y - y^{\delta_n}\|_Y$$

und somit gilt

$$\begin{aligned} \|P_{Q_{\rho_n}} y^{\delta_n} - y\|_Y &\leq \|P_{Q_{\rho_n}} y^{\delta_n} - y^{\delta_n}\|_Y + \|y^{\delta_n} - y\|_Y \\ &\leq \|P_{Q_{\rho_n}} y - y^{\delta_n}\|_Y + \delta_n \\ &\leq \|P_{Q_{\rho_n}} y - y\|_Y + 2\delta_n \end{aligned}$$

Mit (3.4.2) gilt somit $P_{Q_{\rho_n}} y^{\delta_n} \rightarrow y$ für $n \rightarrow \infty$. Mit der selben Argumentation, wie im Satz 3.7, folgt $Au = y$, d.h. $u = u^\dagger$, und die Konvergenz der gesamten Lösungsfolge gegen u^\dagger .

Die Bedingungen sind notwendig, denn $R(A)$ ist nicht abgeschlossen und somit existieren keine inneren Punkte in $R(A)$ (ansonsten wäre das algebraische Innere von $R(A)$ nicht leer. Weiterhin würde die Linearität von A auf den Widerspruch $R(A) = Y$ führen.) Somit kann für jedes $n \in \mathbb{N}$ eine Folge $y^{\delta_n} \in B_{\delta_n}(y) \setminus R(A)$ extrahiert werden. Korollar 3.10 impliziert für entsprechende Lösungen $\|u_{\rho_n}^{\delta_n}\|_U = \rho_n$. Das heißt, wenn $\{\rho_n\}_{n \in \mathbb{N}}$ unbeschränkt ist, so ist $\{u_{\rho_n}^{\delta_n}\}_{n \in \mathbb{N}}$ ebenfalls unbeschränkt und kann dementsprechend nicht schwach*-Konvergent sein (da nach der Folgerung aus dem Banach–Steinhaus Theorem schwach*-konvergente Folgen beschränkt sind). Und mit $u_{\rho_n}^{\delta_n} \rightharpoonup^* u^\dagger$ und der schwach*-Unterhalbstetigkeit der Norm gilt

$$\|u^\dagger\|_U \leq \liminf_{n \rightarrow \infty} \|u_{\rho_n}^{\delta_n}\|_U = \liminf_{n \rightarrow \infty} \rho_n \quad \square$$

Es kann vermerkt werden, dass das Vorliegen zusätzlicher Informationen zu $u^\dagger \in U$ hierbei ausgeschlossen wird. Sicherlich würden solche Informationen zu $u^\dagger \in U$ mit der Wahl $\rho_n \equiv \rho^\dagger$ auf eine konvergente Regularisierung führen. So eine, weder von δ noch von y^δ unabhängige Wahl, wäre jedoch keine Parameterwahlstrategie im Sinne von Definition 2.2.

Mit dem Satz 3.14 werden gewisse Rahmen für das eigentliche Ziel dieses Abschnitts sichtbar. Nun können die klassischen Parameterwahlstrategien s. Definition 2.2 angewandt auf diese Methode diskutiert werden.

A-PRIORI STRATEGIEN Im Grunde impliziert der Satz 3.14 bereits, dass es für eine Datenfolge $y_n \notin R(A)$ keine konvergente a-priori Parameterwahl geben kann, es sei denn, das Problem ist wohl gestellt. Dies wird im nächsten Satz noch explizit hervorgehoben.

Satz 3.15. Sei $\rho_n := \rho(\delta_n)$ auf einer a-priori Parameterwahlstrategie ermittelte Parameterfolge. Wenn für die entsprechende Lösungsfolge gilt $u_{\rho_n}^{\delta_n} \rightharpoonup^* u^\dagger$ für alle $u^\dagger \in U$ und alle $y^{\delta_n} \in B_{\delta_n}(Au^\dagger)$, dann ist die inverse Abbildung A^{-1} stetig.

Beweis. Die Forderung Satz 3.14(ii) muss für jedes $u^\dagger \in U$ gelten. Eine a-priori Parameterwahlstrategie, im Sinne von Definition 2.2, schließt jedoch jegliche Kenntnis über $u^\dagger \in U$ aus. Dies impliziert die Unbeschränktheit der Parameterfolge $\{\rho_n\}_{n \in \mathbb{N}}$ im Widerspruch zu Satz 3.14(i). \square

Im Allgemeinen kann man die vorliegende theoretische Untersuchung des Verfahrens auf einer a-priori Strategie also nicht aufbauen.

HEURISTISCHE STRATEGIEN Dies sind sehr erfolgreiche Strategien, vor allem in Fällen, wenn keine hinreichend scharfe Abschätzung für das Fehlerniveau geliefert werden kann. Unter bestimmten Annahmen existiert eine ganze Klasse von heuristischen Strategien, welche in der Praxis eingesetzt werden s., z.B., [46]; [42]. Auch für endlichdimensionale Inverse Probleme (schlecht konditionierte lineare Gleichungssysteme) sind solche Methoden sehr relevant. Das in der Literatur unter dem Namen Bakushinskii-Veto bekannte Resultat [3] stellt die Anwendung dieser Regel für das oben genannte primäre Ziel dieses Paragraphen jedoch in Frage. Denn sogar im Sinne der klassischen Definition einer Regularisierung [20, Def. 3.1] kann so eine Parameterwahlregel nicht als Teil einer konvergenten Regularisierung für schlecht gestellte Probleme betrachtet werden [20, Thm. 3.3]. Mit dieser Überlegung und Satz 3.15 wird hier weder auf die a-priori, noch auf die heuristische Strategien gesetzt.

A-POSTERIORI STRATEGIEN Eine der bekanntesten a-posteriori Methoden ist sicherlich das Diskrepanzprinzip von Morozov [43], [44]. Mit einem $\tau > 1$ wähle $\rho(\delta, y^\delta)$, so dass

$$(3.4.3) \quad \delta \leq \|Au_{\rho(\delta, y^\delta)}^\delta - y^\delta\|_Y \leq \tau\delta.$$

Nun soll untersucht werden, ob die Bedingungen Satz 3.14 mit der Parameterwahl nach dem Diskrepanzprinzip erfüllt werden.

Satz 3.16. Sei $\rho_n := \rho(\delta_n, y^{\delta_n})$ gewählt nach dem Diskrepanzprinzip (3.4.3) für ein $\tau > 1$. Dann gilt $u_{\rho_n}^{\delta_n} \rightharpoonup^* u^\dagger$ für $n \rightarrow \infty$.

Beweis. Zum Nachweis werden die Eigenschaften der Distanzfunktion s. Kapitel 3.3 genutzt und die Gegebenheit, dass für jedes $0 < \delta_n < \|y^{\delta_n}\|_Y$ nach Satz 3.9 ein ρ_n existiert, so dass gilt (3.4.3). Nun sollen die Bedingungen von Satz 3.14 überprüft werden.

Zuerst soll gezeigt werden, dass $\{\rho_n\}_{n \in \mathbb{N}}$ durch $\|u^\dagger\|_U$ nach oben beschränkt ist. Angenommen es existiert ein $n \in \mathbb{N}$, so dass $\rho_n > \|u^\dagger\|_U$, womit folgt $u^\dagger \in M_{\rho_n}$. Nach Korollar 3.10 ist $\|u_{\rho_n}^{\delta_n}\|_U = \rho_n$ für $\delta_n > 0$, so ist $u_{\rho_n}^{\delta_n} \neq u^\dagger$. Mit der Definition und Eindeutigkeit der Quasilösung folgt

$$\|Au_{\rho(\delta, y^\delta)}^\delta - y^\delta\|_Y = \min_{u \in M_{\rho_n}} \|Au - y^{\delta_n}\|_Y < \|Au^\dagger - y^{\delta_n}\|_Y \leq \delta_n,$$

im Widerspruch zur Parameterwahl ρ_n .

Sei nun angenommen, dass

$$\liminf_{n \rightarrow \infty} \rho_n < \|u^\dagger\|_U =: \rho^\dagger.$$

Da die Ungleichung strikt ist, so existiert eine Teilfolge (mit den selben Indizes bezeichnet), so dass für ein hinreichend kleines $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert mit

$$\rho_n < \tilde{\rho} := \|u^\dagger\|_U - \varepsilon \quad \forall n > N$$

Aus Korollar 3.11 folgt $d(\rho_n, y^{\delta_n}) > d(\tilde{\rho}, y^{\delta_n})$ für alle $n > N$. Mit der Parameterwahlstrategie (3.4.3) und Satz 3.13 folgt dann

$$\lim_{n \rightarrow \infty} \tau\delta_n \geq \lim_{n \rightarrow \infty} d(\rho_n, y^{\delta_n}) \geq \lim_{n \rightarrow \infty} d(\tilde{\rho}, y^{\delta_n}) = d(\tilde{\rho}, y) > 0$$

im Widerspruch zu $\delta_n \rightarrow 0$ (da $u^\dagger \notin M_{\tilde{\rho}}$). □

3 Quasilösungen bei linearen Identifikationsproblemen

Der Nachweis vom Satz 3.16 zeigt im Grunde, dass bei der Wahl der Parameter $\rho_n := \rho(\delta_n, y^{\delta_n})$ nach (3.4.3) gilt

$$\lim_{n \rightarrow \infty} \rho_n = \|u^\dagger\|_U$$

Mit Satz 3.16 steht also fest, dass das Diskrepanzprinzip als Regel zur Parameterwahl, die Hinreichenden Bedingungen vom Satz 3.14 stets erfüllt. Damit ist das eigentliche Ziel erreicht.

3.5 KONVERGENZRATEN

In diesem Abschnitt soll das Verhalten des Regularisierungsfehlers untersucht werden. Doch bedingt durch die Voraussetzungen an den Lösungsraum bzw. die Eigenschaften der Lösungsmenge $M_\rho \subset U$ kann keine starke Konvergenz der Lösungen garantiert werden, so dass keine Aussagen zum Fehler im Sinne von $\|u_\rho^\delta - u^\dagger\|_U$ gemacht werden können. Ein Werkzeug, um in solchen Fällen den Regularisierungsfehler zu erfassen, ist die Charakterisierung mittels Bregmandistanz [10], [30], [47], [22], [25].

Definition 3.17. Für ein konvexes Funktional $J : U \rightarrow \mathbb{R} \cup \{\infty\}$ mit $J(u) < \infty$ ist an der Stelle $u \in U$ ein konvexes Subdifferential gegeben durch

$$(3.5.1) \quad \partial J(u) = \{\xi \in U^* : \langle \xi, v - u \rangle_U \leq J(v) - J(u) \quad \forall v \in U\}.$$

Für $u \in U$ und ein $\xi \in \partial J(u)$ ist die Bregmandistanz gegeben durch

$$(3.5.2) \quad D_J^\xi(v, u) = J(v) - J(u) - \langle \xi, v - u \rangle_U \quad \forall v \in U.$$

Ist J ein konvexes Funktional, so gilt

$$D_J^\xi(v, u) \geq 0 \text{ für jedes } v \in U$$

$$D_J^\xi(u, u) = 0$$

Im Weiteren soll eine äquivalente Problemformulierung zu (3.1.3) betrachtet werden

$$(3.5.3) \quad \min_{u \in U} G(u) + \delta_{M_\rho}(u)$$

mit

$$G(u) := \frac{1}{2} \|Au - y^\delta\|_Y^2$$

so dass die Restriktionen über eine charakteristische Funktion $\delta_M : U \rightarrow \mathbb{R} \cup \{\infty\}$ mit

$$\delta_{M_\rho}(u) := \begin{cases} 0 & \text{wenn } u \in M_\rho \\ \infty & \text{wenn } u \in U \setminus M_\rho \end{cases}$$

gegeben sind.

Um im Rahmen der klassischen Vorgehensweise (für Banachräume vorgestellt in [10]) vorzugehen, müsste $J(u) := \delta_{M_\rho}(u)$ für die Bregmandistanz $D_J^\xi(v, u)$ gewählt werden. Doch so eine Wahl wäre uninformativ für jedes $v \in U \setminus M_\rho$. Daher wird bei der Festlegung der Bregmandistanz das Funktional $J(u)$ wie folgt gesetzt

$$J(u) := \|u\|_U$$

Diese Überlegung basiert auf folgendem Resultat. In der Arbeit [41] wurde gezeigt, dass in topologischen Räumen (U, τ_U) und (Y, τ_Y) die Optimalprobleme, gelöst durch Ivanov und Morozov Regularisierungen, äquivalent sind. Dass heißt

3 Quasilösungen bei linearen Identifikationsproblemen

Wenn eine eindeutige Lösung zu (I) existiert, so ist diese Lösung eindeutig für (M) mit $\|Au_\rho^\delta - y^\delta\| = \delta$.

Wenn eine eindeutige Lösung zu (M) existiert, so ist diese Lösung eindeutig für (I) mit $\|u^\dagger\| = \rho$.

Daher können die Quasilösungen in der Bregmandistanz formal als Lösungen einer Morosov-Regularisierung erfasst werden.

Lemma 3.18. Sei $J(u) := \|u\|_U$ und $\rho = \rho(\delta, y^\delta)$ nach Diskrepanzprinzip (3.4.3) gewählt. Dann gilt folgende Abschätzung

$$(3.5.4) \quad D_J^\xi(u_\rho, u^\dagger) \leq \left| \langle \xi, u_\rho - u^\dagger \rangle_{U^* \times U} \right|$$

Beweis. Nach Satz 3.16 gilt

$$\|u_\rho\| = \rho \leq \|u^\dagger\|$$

und damit $\|u_\rho\| - \|u^\dagger\| \leq 0$ für alle Lösungen gewählt nach (3.4.3). Das heißt, das duale Paar $\langle \xi, u_\rho - u^\dagger \rangle_{U^* \times U}$ kann wegen $D_J^\xi(u, u^\dagger) \geq 0$ ebenfalls nicht positiv sein und es gilt

$$\langle \xi, u_\rho - u^\dagger \rangle_U \leq \|u_\rho\|_U - \|u^\dagger\|_U \leq 0$$

□

Üblicherweise können in Inversen Problemen die Konvergenzraten nur unter zusätzlichen Glattheitsannahmen zu den Daten, sogenannter *Quelldarstellung*, erhalten werden. Im Weiteren wird die klassische Quelldarstellung, wie in [10] vorgeschlagen, betrachtet. Das heißt, es wird folgende Forderung an das Subgradient $\xi \in \partial J(u^\dagger)$ vorausgesetzt

$$(3.5.5) \quad \exists w \in Y^* : \quad \xi = A^* w \in U^*$$

Es ist bekannt, dass im Falle der klassischen Tichonov-Regularisierung mit $J(u) = \|u\|^2$ in einem Hilbertraum U [20] solche Quelldarstellung folgende Form annimmt

$$u^\dagger = A^* w$$

Proposition 3.19. Sei $\rho = \rho(\delta, y^\delta)$ nach Diskrepanzprinzip (3.4.3) gewählt und es existiere ein $w \in Y^*$ mit $\xi := A^* w \in \partial J(u^\dagger)$. Dann existiert ein $C > 0$, so dass gilt

$$(3.5.6) \quad D_J^\xi(u_\rho^\delta; u^\dagger) \leq C\delta.$$

Beweis. Mit den vorausgesetzten Quelldarstellungen und der gewählten Parameterwahlstrategie gilt

$$\begin{aligned} D_J^\xi(u_\rho; u^\dagger) &= \|u_\rho\|_U - \|u^\dagger\|_U - \langle A^* w, u_\rho - u^\dagger \rangle_U \\ &\leq \left| \langle A^* w, u_\rho - u^\dagger \rangle_U \right| = \left| \langle w, A(u_\rho - u^\dagger) \rangle_Y \right| \\ &\leq \|w\|_{Y^*} \|Au_\rho - Au^\dagger\|_Y \\ &\leq \|w\|_{Y^*} (\|Au_\rho - y^\delta\|_Y + \|y^\delta - y\|_Y) \\ &\leq \|w\|_{Y^*} (\tau + 1)\delta. \end{aligned}$$

□

Die Existenz der Quelldarstellung (3.5.5) kann jedoch nicht im Allgemeinen, also nicht für jeden Operator A oder im Falle z.B. einer nicht glatten Funktion $u^\dagger \in U$, garantiert werden. Die Herleitung der Konvergenzraten und vor allem die Antwort nach der Konvergenz in der Bregmandistanz in solchen Fällen bleibt aus. Allerdings könnte z.B. bei der Annahme, dass ein Subgradient für u^\dagger existiert mit $\xi \in L^1(\Omega)$ und der schwach*-Konvergenz der Folge $u_{\rho_n}^{\delta_n} \rightharpoonup^* u^\dagger$ nach Lemma 3.18 behauptet werden

$$D_J^\xi(u_{\rho_n}^{\delta_n}; u^\dagger) \rightarrow 0$$

In [40] ist die Differenzierbarkeit der essential-supremum Norm untersucht worden. Es wurde gezeigt, dass $\partial(\|\cdot\|_{L^\infty(\Omega)})$ nicht Gateaux-Differenzierbar ist. Aus [2] folgt damit, dass für jedes $0 \neq u \in L^\infty(\Omega)$ das Subdifferential nicht aus einem eindeutigen Element besteht. Dabei folgt aus der Konvexität der essential-supremum Norm, dass das Subdifferential nicht leer ist.

Im Kontext der obigen Überlegung wäre die Konstruktion eines Subgradienten $\xi \in L^1(\Omega)$ von Interesse. In einem normierten Raum U besitzt ein Subdifferential über der Norm $\|u\|_U$ folgende Eigenschaften s., z.B., [26]

$$(3.5.7) \quad \begin{aligned} \text{für } u \neq 0 \quad \partial\|\cdot\|_{L^\infty}(u) &= \left\{ \xi \in L^\infty(\Omega)^* : \langle \xi, u \rangle_{L^\infty(\Omega)} = \|u\|_{L^\infty(\Omega)}, \text{ und } \|\xi\|_{L^\infty(\Omega)^*} = 1 \right\} \\ \text{für } u = 0 \quad \partial\|\cdot\|_{L^\infty}(0) &= \left\{ \xi \in L^\infty(\Omega)^* : \|\xi\|_{L^\infty(\Omega)^*} \leq 1 \right\} \end{aligned}$$

Nun soll unter der Geltung dieser Bedingungen ein Subgradient ξ konstruiert werden.

$$\langle \xi, u^\dagger \rangle_{(L^\infty)^* \times L^\infty} = \int_{\Omega} \xi u^\dagger dx = \|u^\dagger\|_{L^\infty(\Omega)} =: \rho^\dagger$$

$$\|\xi\|_{L^1(\Omega)} = \int_{\Omega} |\xi| dx = 1$$

Die Konstruktion ist möglich, wenn der Fall

$$|u^\dagger(x)| < \|u^\dagger\|_{L^\infty(\Omega)} \text{ fast überall}$$

ausgeschlossen ist. Wähle nun $\xi \in L^1(\Omega)$ stückweise konstant, so dass

$$(3.5.8) \quad \xi(x) := \begin{cases} c & \text{wenn } u^\dagger(x) = \rho^\dagger \\ 0 & \text{wenn } |u^\dagger(x)| < \rho^\dagger \\ -c & \text{wenn } u^\dagger(x) = -\rho^\dagger \end{cases}$$

Dann gilt mit $\Omega_0 = \{x \in \Omega : |u^\dagger(x)| = \rho^\dagger\}$

$$\int_{\Omega} \xi u^\dagger dx = c \rho^\dagger \int_{\Omega_0} 1 dx = \rho^\dagger$$

$$\int_{\Omega} |\xi| dx = c \int_{\Omega_0} 1 dx = 1$$

3 Quasilösungen bei linearen Identifikationsproblemen

so dass $c = \left(\int_{\Omega_0} 1 \, dx \right)^{-1}$ die Bedingungen (3.5.7) erfüllt.

So eine Konstruktion des Subgradienten liefert zusammen mit der schwach*-Konvergenz der Lösungen $u_{\rho_n}^{\delta_n} \rightharpoonup^* u^\dagger$ und der Abschätzung aus Lemma 3.18 die Konvergenz der Bregmandistanz

$$(3.5.9) \quad D_j^\xi(u_{\rho_n}^{\delta_n}, u^\dagger) \leq \langle \xi, u_{\rho_n}^{\delta_n} - u^\dagger \rangle_{U^* \times U} \rightarrow 0 \quad \text{für } n \rightarrow \infty$$

Ein Werkzeug, um Konvergenzraten herzuleiten fehlt hier jedoch, da die im (3.5.8) vorgeschlagene Variante zur Konstruktion eines Subgradienten die Quelldarstellung (3.5.5) nicht notwendigerweise erfüllt.

3.6 ANWENDUNG: NUMERISCHE LÖSUNG, LINEARER FALL

In diesem Kapitel sollen die theoretischen Erkenntnisse in die Praxis umgesetzt werden. Seien die Funktionenräume dazu wie folgt gewählt $U = L^\infty(\Omega)$ und $Y = L^2(\Omega)$ für ein beschränktes Lipschitz-Gebiet $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$. Das primäre Ziel dieses Abschnitts liegt im Vergleich von theoretischen Resultaten und der realen Möglichkeiten der untersuchten Regularisierungsmethode in nicht-reflexiven Banachräumen. Der Anwendung soll daher zunächst ein möglichst einfaches Modell dienen. Zu Modellierung des linearen Problems (3.1.3) wird daher die Rekonstruktion der Temperaturquelle $u(x) \in L^\infty(\Omega)$ zum beobachteten Zustand $y \in H^1(\Omega)$ gewählt, welche durch eine elliptische Partielle Differentialgleichung mit homogenen Neumann Randbedingungen beschrieben ist

$$(3.6.1) \quad \begin{cases} -\Delta y + cy = u & \text{in } \Omega, \\ \partial_\nu y = 0 & \text{auf } \partial\Omega, \end{cases}$$

für ein gegebenes $c > 0$.

$$A : L^\infty(\Omega) \rightarrow L^2(\Omega), \quad u \mapsto y,$$

entspricht den in Annahme 3.1 geforderten Bedingungen an die Abbildung linear, injektiv, stetig, und schwach*-abgeschlossen zu sein. Es ist bekannt, dass u_ρ in so einem Fall Lösung des Problems (3.1.3) ist (welche in diesem Kontext als *bang-bang-control Problem* bekannt ist s., z.B., [18]) genau dann, wenn

$$(3.6.2) \quad u_\rho = \text{proj}_{M_\rho}(u_\rho - p_\rho),$$

wo

$$\text{proj}_{M_\rho} : L^2(\Omega) \rightarrow L^2(\Omega), \quad [\text{proj}_{M_\rho}(v)](x) =: \begin{cases} \rho & \text{if } v(x) > \rho \\ v(x) & \text{if } |v(x)| \leq \rho \\ -\rho & \text{if } v(x) < -\rho \end{cases}$$

und p_ρ ist die Lösung zum *adjungierten Problem*

$$(3.6.3) \quad \begin{cases} -\Delta p + cp = f & \text{in } \Omega, \\ \partial_\nu p = 0 & \text{auf } \partial\Omega, \end{cases}$$

für $f = Au_\rho - y^\delta$; s., z.B., [55, Lem. 2.26]. Hier wird davon ausgegangen, dass der adjungierte Zustand $p_\rho \in H^1(\Omega)$ die punktweise fast überall Auswertung zulässt.

Unter der Nutzung der Semi-Smooth-Newton Methode [12], [13], [27], [56] mit der lokal-superlinearen Konvergenz soll (3.6.2) gelöst werden. Es ist aber bekannt, dass die punktweise Projektion von $L^p(\Omega) \rightarrow L^q(\Omega)$ genau dann semi-smooth ist, wenn $p > q$ gilt. Daher ist (3.6.2) nicht unbedingt semi-smooth im Bezug zu u_ρ . Gewöhnlich wird dann zum Funktional (3.1.3) ein stabilisierender Strafterm addiert. Dies erlaubt u_ρ innerhalb der Projektion zu eliminieren, s., z.B., [27], [28] oder auch [55, Thm. 2.28, Ch. 2.12.4]. Da der Schwerpunkt dieser Arbeit der Ivanov-Regularisierung gilt, wird dieser Weg hier nicht verfolgt und stattdessen eine vorherige Diskretisierung vorgenommen.

3 Quasilösungen bei linearen Identifikationsproblemen

Sei $Y_h \subset L^2(\Omega)$ ein endlich dimensionaler Raum, aufgespannt durch punktweise lineare Basis-Funktionen (Triangulation bei der Finite-Elemente-Methode.) Die Funktionen $y_h, u_h, p_h \in Y_h$ werden nun mit dem Vektor ihrer Basiskoeffizienten identifiziert, welche in diesem Fall mit den Werten in den Knoten übereinstimmen (bzw. ihr Maximum und Minimum in den Knoten erreichen).

Durch die Einführung der "Stiffness"-Matrix $K_h \in \mathbb{R}^{N \times N}$ und der "Mass"-Matrix $M_h \in \mathbb{R}^{N \times N}$ werden die diskretisierten Bedingungen

$$(3.6.4) \quad \begin{cases} (K_h + cM_h)y_h = M_h u_h, \\ (K_h + cM_h)p_h = M_h(y_h - y_h^\delta), \\ u_h = \text{proj}_{[-\rho, \rho]}(u_h - p_h), \end{cases}$$

formuliert, wo $y_h^\delta \in Y_h$ die L^2 Projektion auf y^δ bezeichnet, so dass diese Projektion nun komponentenweise gemeint ist und semi-smooth von \mathbb{R}^N nach \mathbb{R}^N ist, wo Newton-Ableitung in $v \in \mathbb{R}^N$ in Richtung $w \in \mathbb{R}^N$ komponentenweise gegeben ist durch

$$(3.6.5) \quad [D_N \text{proj}_{[-\rho, \rho]}(v)w]_i = \begin{cases} w_i & \text{wenn } |w_i| \leq \rho, \\ 0 & \text{sonst} \end{cases}$$

Die aktiven and inaktiven Mengen sind gegeben durch

$$(3.6.6) \quad \begin{cases} \mathcal{A}_+ := \{i : u_i - p_i > \rho\} \\ \mathcal{A}_- := \{i : u_i - p_i < -\rho\} \\ \mathcal{I} := \{i : |u_i - p_i| \leq \rho\} \end{cases}$$

und die entsprechenden Indikatorfunktionen $\mathbb{1}_{\mathcal{I}}, \mathbb{1}_{\mathcal{A}_+}, \mathbb{1}_{\mathcal{A}_-}$ durch

$$\mathbb{1}_{\mathcal{A}} = \begin{cases} 1 & \text{für } i \in \mathcal{A}_{+/-} \\ 0 & \text{sonst} \end{cases} \quad \mathbb{1}_{\mathcal{I}} = \begin{cases} 1 & \text{für } i \notin \mathcal{A}_{+/-} \\ 0 & \text{sonst} \end{cases}$$

Des Weiteren seien $O \in \mathbb{R}^{N \times N}$ eine Null-Matrix und $D_{\mathcal{I}} \in \mathbb{R}^{N \times N}, D_{\mathcal{A}_{+/-}} \in \mathbb{R}^{N \times N}$ Diagonal-Matrizen mit den Einträgen, gegeben durch $\mathbb{1}_{\mathcal{I}}, \mathbb{1}_{\mathcal{A}_+}, \mathbb{1}_{\mathcal{A}_-}$. Mit $y_h^{k+1} = y_h^k + \delta y, p_h^{k+1} = p_h^k + \delta p,$ und $u_h^{k+1} = u_h^k + \delta u$ kann nun ein Newton-Schritt wie folgt beschrieben werden

$$(3.6.7) \quad B_k \begin{pmatrix} \delta y \\ \delta p \\ \delta u \end{pmatrix} := \begin{pmatrix} K_h + cM_h & O & -M_h \\ -M_h & K_h + cM_h & O \\ O & D_{\mathcal{I}} & D_{\mathcal{A}_+} + D_{\mathcal{A}_-} \end{pmatrix} \begin{pmatrix} \delta y \\ \delta p \\ \delta u \end{pmatrix} \\ = - \begin{pmatrix} (K_h + cM_h)y_h^k - M_h u_h^k \\ (K_h + cM_h)p_h^k - M_h(y_h^k - y_h^\delta) \\ u_h - \text{proj}_{[-\rho, \rho]}(u_h^k - p_h^k) \end{pmatrix} =: -b^k$$

Der Iterationsprozess wird beendet, wenn entweder eine feste (vorher heuristisch festgelegte) Anzahl von Iterationen überschritten wird oder die aktiven Mengen $\mathcal{A}_+^k = \mathcal{A}_+^{k-1}$

und $\mathcal{A}_-^k = \mathcal{A}_-^{k-1}$ ergeben. Im letzteren Fall ist (p_h^k, y_h^k, u_h^k) bereits eine Lösung zu (3.6.4); s. [32, Rem. 7.1.1]. Um eine gewisse Robustheit des Verfahrens zu erreichen, soll eine Dämpfung $\|b^k\|_2 \leq \text{tol}$ eingeführt werden mit einem ebenfalls heuristisch gewählten $\text{tol} > 0$ um einen gelungenen Ausstieg aus der Schleife zu erreichen. Die vollständige Prozedur ist im Algorithmus 1 beschrieben.

Algorithm 1 Damped semi-smooth Newton method

Input: $y_h^\delta, \rho > 0, q \in (0, 1), i_{\max}, k_{\max}, \text{tol} > 0, y_h^0, p_h^0, u_h^0$

- 1: **for** $k = 1, \dots, k_{\max}$ **do**
- 2: compute active sets $\mathcal{A}_+^k, \mathcal{A}_-^k, \mathcal{I}^k$ from (3.6.6)
- 3: compute Newton matrix B_k , right-hand side b^k from (3.6.7)
- 4: **if** $\|b_k\|_2 < \text{tol}$ **then**
- 5: set $\text{conv} = \text{true}$; **return**
- 6: solve Newton step (3.6.7) for $(\delta y, \delta p, \delta u)$
- 7: **for** $i = 1, \dots, i_{\max}$ **do**
- 8: set $y_h^i + q^{i-1}\delta y, p_h^i + q^{i-1}\delta p, u_h^i + q^{i-1}\delta u$
- 9: compute b^i from (3.6.7)
- 10: **if** $\|b^i\|_h < \|b^k\|_h$ **then**
- 11: set $y_h^{k+1} = y_h^i, p_h^{k+1} = p_h^i, u_h^{k+1} = u_h^i$; **break**
- 12: set $\text{conv} = \text{false}$

Output: $y_h^k, p_h^k, u_h^k, \text{conv}$

Zur Ermittlung der optimalen Parameter, welche dem Diskrepanzprinzip genügen, wäre die Anwendung der Bisektionsmethode naheliegend. In der Praxis muss jedoch damit gerechnet werden, dass die Newton-Methoden nur die lokale Konvergenz garantieren. Die vorgeschlagene gedämpfte Newton-Methode zeigt zwar fast sichere globale Konvergenzeigenschaften, führt aber oft zum enormen Rechenaufwand wegen großer Iterationsanzahl. Daher wird hier auf eine Art "Backtracking-Methode" gesetzt. Dieser Ansatz basiert auf folgender Beobachtung

- Es existiert ein hinreichend großes $\hat{\rho}$, so dass das Semi-Smooth-Newton Verfahren in einem Schritt konvergiert. Dies wird dadurch hervorgerufen, dass die Abbildung $A_h := (K_h + cM_h)^{-1}$ aufgrund der Diskretisierung invertierbar ist und jedes y_h^δ im Bild des Operators A_h liegt.
- Die ermittelte Lösung für ein ρ_k ist ein guter Startpunkt für $\rho_{k+1} < \rho_k$, wenn die Differenz $|\rho_{k+1} - \rho_k|$ nicht zu groß ist. Dies wird dadurch hervorgerufen, dass zum einen der "Downstep" $\rho_{k+1} < \rho_k$ beibehalten werden muss, damit die bereits ermittelte Information über die aktiven Mengen $\mathcal{A}_k := \{\mathcal{A}_+ \cup \mathcal{A}_-\}_k$ in jeweils nächstem Schritt nicht komplett aufs Neue ermittelt werden muss, denn für $\rho_{k+1} > \rho_k$ werden die aktiven Mengen \mathcal{A}_{k+1} auf Null gesetzt. Zum anderen diktiert die Schrittgröße $|\rho_{k+1} - \rho_k|$ den Level der Veränderung von aktiven Mengen.

Ob ein nächster Schritt ρ_{k+1} passend gesetzt wurde, wird über eine fixierte Anzahl N der Iterationen einer Newton-Schleife entschieden. Ist die Iterationsanzahl überschritten, so ist

3 Quasilösungen bei linearen Identifikationsproblemen

der Schritt $|\rho_{k+1} - \rho_k|$ zu groß gewählt worden und muss verkleinert werden. Die sichere lokale Konvergenz der angewendeten Newton-Methode und die oben beschriebene Beobachtungen motivieren eine "Backtracking"-Vorgehensweise, anstatt der reinen Bisektionsmethode. Diese wird in drei Phasen aufgeteilt.

- I. Starten mit einem beliebigen $\rho \geq 0$. Zunächst wird ρ vergrößert solange, bis ein ρ gefunden wird für welches die Newton-Methode in $n \leq N$ Schritten konvergiert und die Diskrepanz kleiner als das Fehlerniveau δ ist.
- II. Dann wird ρ stark herunter gesetzt solange, bis die Newton-Methode in einem ρ_k nicht mehr konvergiert (in $n \leq N$ Schritten).
- III. Gestartet vom vorletzten Punkt ρ_{k-1} , unter der Beachtung der Konvergenz der Newton-Methode (in $n \leq N$ Schritten), wird nun das optimale ρ gesucht, welches das Diskrepanzprinzip (3.4.3) erfüllt.

Die Phasen I und II spielen eine Rolle der groben Vorbereitung, ergeben aber eine schnelle Annäherung an das optimale ρ . Natürlich beinhalten auch diese Phasen das Ausstiegskriterium für den Fall, dass das Diskrepanzprinzip bereits erfüllt ist. Die vollständige Prozedur zur Ermittlung des optimalen Parameters ist im Algorithmus 2 gegeben.

Nun soll die praktische Anwendbarkeit der Ivanov-Regularisierung anhand eines numerischen Beispiels am Testproblem (3.6.1) betrachtet werden. Dazu wird der N -dimensionale Funktionenraum auf der Platte $\Omega = [-1, 1]^2 \subset \mathbb{R}^2$ beim Diskretisieren durch einheitliche Friedrichs-Keller-Triangulation gegeben, wo $N = 128 \times 128$ Eckpunkte sind. Der Potenzialkoeffizient sei mit $c = 1$ festgelegt. Die zu rekonstruierende Funktion u^\dagger wird durch unterschiedlich ausgerichtete (positive und negative) Säulen mit unterschiedlicher Höhe gegeben, wobei $\rho^\dagger := \|u^\dagger\|_{L^\infty(\Omega)} = 4$ s., Abb.3.2a. Die gestörten Daten werden simuliert als

$$y^\delta = y^\dagger + \frac{s}{100} \|y^\dagger\|_\infty \frac{\eta}{\|\eta\|_2},$$

wo $y^\dagger = A_h u^\dagger$, $s > 0$ die relative Störung gegeben in Prozent, $\eta \in \mathbb{R}^N$ ein Zufallsvektor (random vector) mit η_i , $1 \leq i \leq N$ unabhängig normal verteilt mit Mittelwert 0 und Standardabweichung 1, und $\|\eta\|_2 := \eta^T M_h \eta$ die diskrete $L^2(\Omega)$ -Norm ist. Die in Algorithmen 1 und 2 verwendeten Parameter werden wie folgt festgelegt

- Im Algorithmus 1 seien $q = 0.7$, $i_{\max} = 10$, $k_{\max} = 30$ und $\text{tol} = 10^{-9}$.
- Im Algorithmus 2 seien $\tau = 1.1$ and $\rho_0 = 10$.

Unter Anwendung von Matlab werden nun die in den Algorithmen 1 und 2 gegebenen Verfahren implementiert.

Die Abbildungen 3.2 zeigen die Lösungen zu $s = 1, 0.02, 0.01, 0.001, 0.0001$ mit den entsprechenden Parametern $\rho = \rho(\delta, y^\delta) \approx 2.5, 3.896, 3.952, 3.995, 3.999$ welche sich nach Diskrepanzprinzip ergeben. In allen Fällen gilt dabei, wie auch stützend auf das theoretische Resultat in 3.10 erwartet, dass $\|u_\rho^\delta\|_{L^\infty(\Omega)} = \rho$. Bei sinkender Störung s , und damit auch Diskrepanz

Algorithm 2 Parameter choice

Input: $y^\delta, \delta, \tau > 1, \rho_0 \geq 0$

- 1: set $y_h^k = 0, p_h^k = 0, u_h^k = 0$
- 2: **for** $k = 0, \dots$ **do** ▷ Phase I
- 3: compute y_h^k, p_h^k, u_h^k , conv using Algorithm 1 for $\rho_k, y_h^{k-1}, p_h^{k-1}, u_h^{k-1}$
- 4: compute $d_k = \|y_k - y^\delta\|_2$
- 5: **if** $d_k < \delta$ and conv = true **then**
- 6: **break**
- 7: set $\rho_{k+1} = \rho_k + \rho_0$
- 8: set $\rho_0 = \rho_k/2, y_h^0 = y_h^k, p_h^0 = p_h^k, u_h^0 = u_h^k$
- 9: **for** $k = 0, \dots$ **do** ▷ Phase II
- 10: compute y_h^k, p_h^k, u_h^k , conv using Algorithm 1 for $\rho_k, y_h^{k-1}, p_h^{k-1}, u_h^{k-1}$
- 11: compute $d_k = \|y_k - y^\delta\|_2$
- 12: **if** $d_k > \delta$ or conv = false **then**
- 13: **break**
- 14: set $\rho_{k+1} = \rho_k/2$
- 15: set $\Delta\rho = \rho_{k-1}/2, \rho_0 = \rho_{k-1} - \Delta\rho, y_h^0 = y_h^k, p_h^0 = p_h^k, u_h^0 = u_h^k$
- 16: **for** $k = 0, \dots$ **do** ▷ Phase III
- 17: compute y_h^k, p_h^k, u_h^k , conv using Algorithm 1 for $\rho_k, y_h^{k-1}, p_h^{k-1}, u_h^{k-1}$
- 18: compute $d_k = \|y_k - y^\delta\|_Y$
- 19: **if** $\delta \leq d_k \leq \tau\delta$ and conv = true **then**
- 20: **break**
- 21: **else if** $d_k > \tau\delta$ or conv = false **then**
- 22: set $\Delta\rho \leftarrow \Delta\rho/2$
- 23: set $\rho_{k+1} = \rho_k + \Delta\rho$
- 24: **else**
- 25: set $\rho_{k+1} = \rho_k - \Delta\rho$

Output: $y_h^k, p_h^k, u_h^k, \rho_k$

$\|y^\delta - y^\dagger\|_2$, steigt die Qualität, bis die Lösung praktisch identisch mit u^\dagger ist (in diesem Experiment für $s = 0.0001$). Im Grunde ist die Rekonstruktion bereits für $s \approx 0.01$ akzeptabel. Hier kann vermerkt werden, dass sogar für große s die Form und vor allem die Lokation der Säulen gut wiederhergestellt werden. Dabei zeigen die Abbildungen keine Glättung, wie es z.B. im Falle einer Tichonov-Regularisierung zu erwarten wäre. Für die größten Säulen mit $u_\rho^\dagger(x) = \rho^\dagger$ ist die Form bereits für $s = 1$ gut erkennbar.

Die Tabelle 3.1 zeigt folgende Daten zu jedem prozentual gegebenem Fehler s an

- die tatsächliche Störung δ
- die Diskrepanz $d(\rho, y^\delta)$ zum entsprechenden optimalen Parameter $\rho(\delta, y^\delta)$, welche durch Algorithmus 2 gegeben werden
- den entsprechenden Rekonstruktionsfehler der Quasilösung, gegeben in $L^\infty(\Omega)$ -Norm,

3 Quasilösungen bei linearen Identifikationsproblemen

$L^2(\Omega)$ -Norm und durch das duale Paar $\langle \xi, u_\rho^\delta - u^\dagger \rangle$ mit $\xi \in L^1(\Omega) \subset L^\infty(\Omega)^*$.

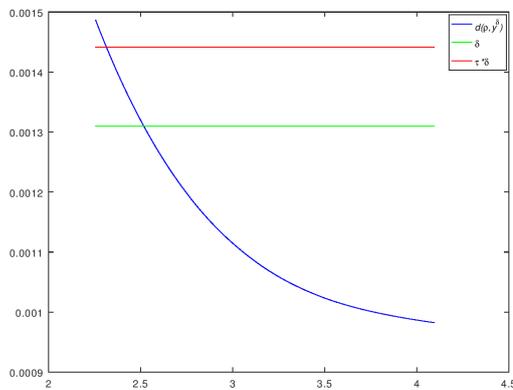
Zum Schluss dieses Kapitels sollen die beobachteten Resultate kurz zusammengefasst werden. Dabei verdient das Konvergenzverhalten der Lösungen, gegeben durch den Ausdruck $\langle \xi, u_{\rho_n}^{\delta_n} - u^\dagger \rangle_{U^* \times U}$ ein besonderes Interesse. Die Überlegung den Rekonstruktionsfehler zum Testproblem durch das duale Paar zu erfassen (s. Tabelle 3.1), basiert auf der Konstruktion des Subgradienten $\xi \in L^1(\Omega) \subset L^\infty(\Omega)^*$ (3.5.8) im Kapitel 3.5. Diese Konstruktion liefert zusammen mit der schwach*-Konvergenz der Lösungen $u_{\rho_n}^{\delta_n} \rightharpoonup^* u^\dagger$ und der Abschätzung aus Lemma 3.18 die Konvergenz der Bregmandistanz (3.5.9)

$$(3.6.8) \quad D_J^\xi(u_{\rho_n}^{\delta_n}, u^\dagger) \rightarrow 0 \quad \text{für } n \rightarrow \infty$$

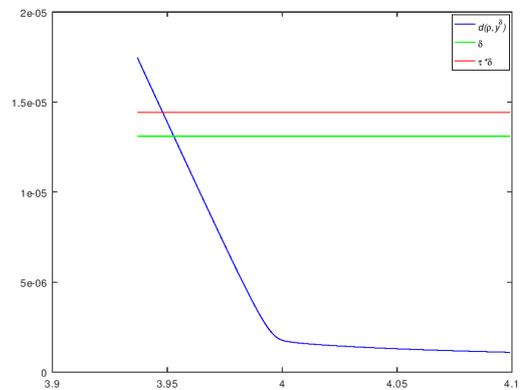
Die Daten aus der Tabelle ergeben, dass diese Konvergenz sich wie $O(\delta)$ verhält. Allerdings gibt es keine theoretischen Anhaltspunkte um diese Konvergenzraten etablieren zu können.

Der Vollständigkeit halber wird in der Tabelle 3.1 der Fehler zusätzlich sowohl in $L^2(\Omega)$ -Norm $\|u_\rho^\delta - u^\dagger\|_2$, als auch in $L^\infty(\Omega)$ -Norm $\|u_\rho^\delta - u^\dagger\|_\infty$ angegeben. Hier im diskreten Fall, liefert diese starke Konvergenz jedoch keinen Anhaltspunkt vom theoretischen Standpunkt her.

Die mittels Matlab berechneten Diagramme 3.1a und 3.1b zeigen das im Kapitel 3.3 diskutierte Verhalten der Distanzfunktion für gegebene Daten (δ, y^δ) . Die zur numerischen Lösung verwendete Semi-Smooth-Newton Methode liefert durchaus brauchbare Rekonstruktionen der Lösungen u_ρ^δ . Anhand der Abbildungen 3.2 kann man erkennen, dass für sehr große Störungen die Rekonstruktion eher einer Bang-Bang-Steuerung ähnelt. Im Allgemeinen verhält sich die Regularisierung beim Anwenden von Diskrepanzprinzip so, wie gestützt auf die theoretischen Resultate auch erwartet.

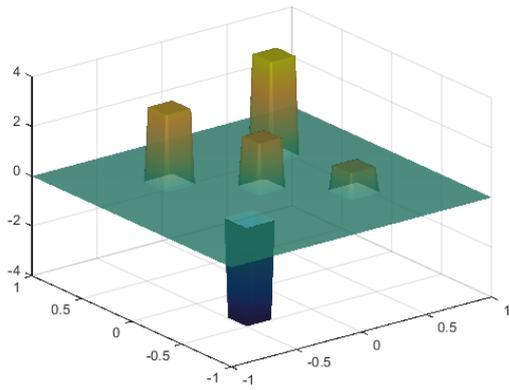


(a) $\rho \approx 2.5$ für $s = 1\%$

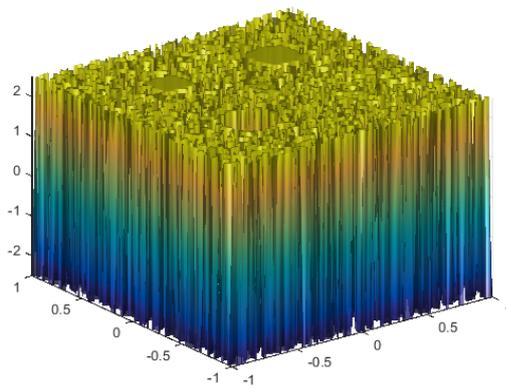


(b) $\rho \approx 3.95$ für $s = 0.01\%$

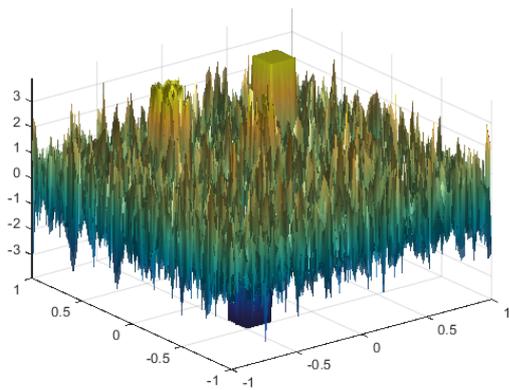
Abbildung 3.1: Distanzfunktion für unterschiedliche Störungslevel



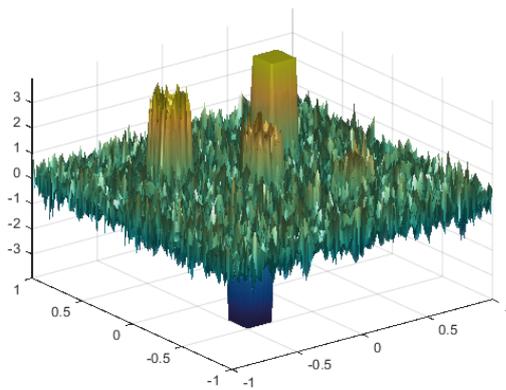
(a) u^\dagger mit $\|u^\dagger\|_{L^\infty(\Omega)} = 4$



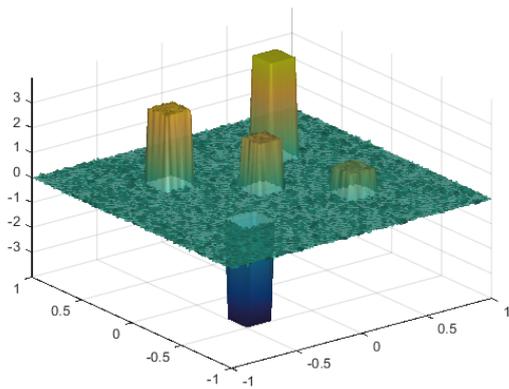
(b) u_ρ^δ für $s = 1\%$, $\rho \approx 2.5$



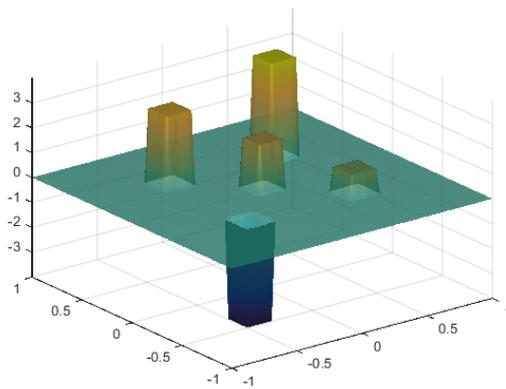
(c) u_ρ^δ für $s = 0.02\%$, $\rho \approx 3.896$



(d) u_ρ^δ für $s = 0.01\%$, $\rho \approx 3.952$



(e) u_ρ^δ für $s = 0.001\%$, $\rho \approx 3.995$



(f) u_ρ^δ für $s = 0.0001\%$, $\rho \approx 3.999$

Abbildung 3.2: Die exakten Daten u^\dagger und Quasilösungen u_ρ^δ für unterschiedliche Störungslevel

Tabelle 3.1: Quantitative Resultate

s	δ	$d(\rho, y^\delta)$	$\rho(\delta, y^\delta)$	$\ u_\rho^\delta - u^\delta\ _\infty$	$\ u_\rho^\delta - u^\delta\ _2$	$\langle \xi, u_\rho^\delta - u^\delta \rangle$
10^0	$1.310 \cdot 10^{-3}$	$1.321 \cdot 10^{-3}$	2.500	$4.500 \cdot 10^0$	$3.571 \cdot 10^0$	$1.500 \cdot 10^0$
10^{-1}	$1.310 \cdot 10^{-4}$	$1.400 \cdot 10^{-4}$	3.594	$5.594 \cdot 10^0$	$3.101 \cdot 10^0$	$4.063 \cdot 10^{-1}$
10^{-2}	$1.310 \cdot 10^{-5}$	$1.358 \cdot 10^{-5}$	3.952	$2.894 \cdot 10^0$	$5.620 \cdot 10^{-1}$	$4.836 \cdot 10^{-2}$
10^{-3}	$1.310 \cdot 10^{-6}$	$1.367 \cdot 10^{-6}$	3.995	$2.051 \cdot 10^{-1}$	$5.661 \cdot 10^{-2}$	$4.944 \cdot 10^{-3}$
10^{-4}	$1.310 \cdot 10^{-7}$	$1.391 \cdot 10^{-7}$	3.999	$2.317 \cdot 10^{-2}$	$5.726 \cdot 10^{-3}$	$4.987 \cdot 10^{-4}$
10^{-5}	$1.310 \cdot 10^{-8}$	$1.357 \cdot 10^{-8}$	4.000	$2.070 \cdot 10^{-3}$	$5.612 \cdot 10^{-4}$	$4.844 \cdot 10^{-5}$

4 QUASILÖSUNGEN BEI NICHTLINEAREN IDENTIFIKATIONSPROBLEMEN

4.1 EXISTENZ UND STABILITÄT

In diesem Kapitel werden Identifikationsprobleme mit einem nicht-linearen Operator untersucht. Es wird ein direktes Problem betrachtet, gegeben durch die Operatorgleichung

$$(4.1.1) \quad F(u) = y$$

mit einem stetigen Operator $F : U \rightarrow Y$. Die Funktionenräume U, Y sind wie in 3.1 vorausgesetzt

- Y ein gleichmäßig konvexer Banachraum
- $U \equiv L^\infty(\Omega)$, und $\Omega \subset \mathbb{R}^N$ beschränktes Gebiet

Wie auch im linearen Fall wird angenommen, dass die exakten Daten $y \in R(F) \subset Y$ im Bild des Operators liegen. Vorhanden sind jedoch ausschließlich durch ein Rauschen gestörte Daten $y^\delta \in Y$. Weiterhin wird angenommen, dass der Störungslevel $\delta \geq 0$ angegeben werden kann, so dass gilt

$$(4.1.2) \quad \|y^\delta - y\|_Y \leq \delta$$

Die Ivanov-Regularisierung führt im nicht-linearen Fall auf folgendes Optimierungsproblem

$$(4.1.3) \quad \inf_{u \in M_\rho} \|F(u) - y\|_Y \quad M_\rho := \{u \in L^\infty(\Omega) : \|u - u^*\|_{L^\infty} \leq \rho\}$$

wo die Menge M_ρ zusätzlich über ein *Referenzelement* $u^* \in U$ gegeben ist, und

$$(4.1.4) \quad d(\rho, y) := \inf_{u \in M_\rho} \|F(u) - y\|_Y$$

nach Definition die Distanzfunktion ist (s. (3.3.1)).

Im Unterschied zum linearen Fall muss eine stärkere Forderung an die Abgeschlossenheit der Abbildung gemacht werden.

Annahme 4.1. Sei die stetige Abbildung $F : U \rightarrow Y$ schwach*-schwach-abgeschlossen, das heißt für jede Folge $\{u_n\}$ mit $u_n \rightharpoonup^* u$ in U und $F(u_n) \rightarrow y$ in Y folgt $y = F(u)$.

- Eine Forderung, der Operator F sei schwach*-schwach-stetig, wäre für die vorliegende Untersuchung zu stark. Dazu kann eine Bildfolge $\{F(u_n)\}_{n \in \mathbb{N}}$ zu der schwach*-konvergenten Lösungsfolge $u_n \rightharpoonup^* u_\rho$ betrachtet werden. Im Unterschied zu (3.1.4) im Abschnitt 3.1 ist die Bildmenge

$$Q_\rho := \{F(u) \in Y : u \in M_\rho\}$$

nicht notwendigerweise beschränkt für ein gegebenes $\rho \geq 0$, denn F ist stetig aber nicht linear. Falls es jedoch eine Schranke zu der Bildfolge $\{F(u_n)\}$ gibt, was z.B. bei bestimmten Parameterwahlbedingungen der Fall ist, so existiert eine schwach-konvergente Teilfolge, so dass mit der schwach*-schwach Abgeschlossenheit des Operators und der Teilfolgen-Teilfolgen Argumentation die schwache Konvergenz für die ganze Bildfolge $F(u_n) \rightharpoonup F(u)$ gilt.

Annahme 4.2. *Es existiert $\rho > 0$, so dass die Lösung der Operatorgleichung 4.1.1 eindeutig in der Umgebung $B_\rho(u^*)$ ist. Das heißt die u^* -Minimum-Norm-Lösung ist zu den exakten Daten $y = F(u^\dagger)$ eindeutig.*

- Die Forderung nach der Injektivität der Abbildung in Annahme 3.1 wäre hier im nichtlinearen Fall nicht sinngemäß. Falls eine Lösung u_ρ zum Problem (4.1.3) existiert, so kann mit dem Konzept der u^* -Minimum-Norm-Lösung

$$(4.1.5) \quad \|u_\rho - u^*\|_U = \min \left\{ \|\hat{u} - u^*\|_U : \|F(\hat{u}) - y\|_Y = \min_{u \in M_\rho} \|F(u) - y\|_Y \right\}$$

in günstigen Fällen die lokale Eindeutigkeit direkt erreicht werden [29]. Sonst muss diese Eindeutigkeit eben gefordert werden. Im betrachteten Fall wäre so eine Forderung jedoch zu stark, da der Lösungsraum nicht strikt konvex ist. Deswegen wird hier die Eindeutigkeit der u^* -Minimum-Norm-Lösung in der Operatorgleichung (4.1.1) und ausschließlich für die exakten Daten $y = F(u^\dagger)$ gefordert. Denn ohne diese Forderung wäre zusätzlich zu der *Stabilitätsbedingung* auch noch die *Eindeutigkeitsbedingung* der Hadamardschen Definition zur Korrektheit eines Problems nicht erfüllt.

Satz 4.3. *Für jedes $y \in Y$ und jedes $\rho \geq 0$ existiert mindestens eine Lösung $u_\rho \in M_\rho$ zum Problem (4.1.3).*

Beweis. Wie auch in Satz 3.4 ist $G : Y \supset Q_\rho \rightarrow \mathbb{R}$, $G(q) := \|q - y\|_Y$ ein stetiges beschränktes Funktional mit der nicht-negativen unteren Schranke

$$0 \leq \inf \{ \|q - y\|_Y : q \in Q_\rho \} = r$$

Die Bildmenge Q_ρ ist im Allgemeinen weder beschränkt noch konvex. Daher wird eine Schnittmenge

$$S := Q_\rho \cap B_{r+1}(y)$$

betrachtet, welche nicht leer, beschränkt und mit der Annahme 4.1 schwach-abgeschlossen ist. Damit existiert nach Satz 3.4 ein Minimierer

$$q_\rho = \min \{ \|q - y\|_Y : q \in S \}$$

Offensichtlich ist q_ρ ebenfalls ein Minimierer auf der Menge Q_ρ , so dass alle $\{u_\rho : F(u_\rho) = q_\rho\}$ Lösungen des Problems sind. \square

Im Unterschied zum linearen Fall, sind die Lösungen nicht eindeutig bestimmt. Sogar die Selektion nach u^* -Minimum-Norm-Lösung hilft hier nicht weiter, da der Lösungsraum nicht strikt konvex ist. So können mehrere Lösungen $\{u_\rho\}$ mit $\|u^* - u_\rho\|_U = \rho$ existieren. Diese Überlegung wird im Nachweis zur Stabilität ebenfalls verankert.

Einige Eigenschaften der Distanzfunktion, welche bereits im Kapitel 3.3 gezeigt wurden, sind auf den nicht-linearen Fall übertragbar und werden zum Nachweis der Behauptungen auch in diesem Kapitel von Bedeutung sein. So zum Beispiel die Lipschitzstetigkeit der Distanzfunktion in der Variablen $y \in Y$ s. Satz 3.13.

Lemma 4.4. *Sei die Voraussetzung (4.1.2) für $y^\delta \in Y$ und die entsprechenden exakten Daten $y = F(u^\dagger)$ erfüllt, d.h. $\|y^\delta - y\|_Y \leq \delta$. Sei durch $\rho^\dagger := \|u^* - u^\dagger\|$ der entsprechende Parameter zur eindeutigen lokalen Lösung s . Annahme 4.2 gekennzeichnet. Dann ist die Distanzfunktion Lipschitzstetig in der Variablen $y \in Y$ und es gilt folgende Ungleichung*

$$(4.1.6) \quad d(\rho^\dagger, y^\delta) \leq \delta$$

Beweis. Die Distanzfunktion $d(\rho, y)$ ist für jedes feste $\rho \geq 0$ Lipschitzstetig mit Konstante 1 in der Variablen $y \in Y$. Seien dazu die Repräsentanten der Menge der entsprechenden Lösungen von $y_1 := y^{\delta_1}$ und $y_2 := y^{\delta_2}$ mit $\|y^{\delta_1}\| \neq \|y^{\delta_2}\|$ durch u_1 und u_2 gekennzeichnet mit $q_1 = F(u_1) = P_{Q_\rho} y_1$ und $q_2 = F(u_2) = P_{Q_\rho} y_2$. Es gilt

$$\begin{aligned} d(\rho, y_1) &= \|q_1 - y_1\|_Y \leq \|q_2 - y_1\|_Y \leq \|q_2 - y_2\|_Y + \|y_2 - y_1\|_Y \\ d(\rho, y_2) &= \|q_2 - y_2\|_Y \leq \|q_1 - y_2\|_Y \leq \|q_1 - y_1\|_Y + \|y_1 - y_2\|_Y \end{aligned}$$

$$\text{und damit} \quad |d(\rho, y_1) - d(\rho, y_2)| \leq \|y_1 - y_2\|_Y$$

Weiterhin ist für $y = F(u^\dagger)$ nach der Definition der Distanzfunktion $d(\rho^\dagger, y) = 0$, so dass folgende Ungleichung gilt

$$d(\rho^\dagger, y^\delta) - d(\rho^\dagger, y) \leq \|y^\delta - y\| \leq \delta$$

\square

Bemerkung 4.5. *Es ist wichtig zu vermerken, dass die Konvexität der Bildmenge Q_ρ in der Beweisführung von Lemma 4.4, wie auch im Satz 3.13, nicht verwendet wurde. Der Nachweis der Lipschitzstetigkeit der Distanzfunktion in ihrem zweiten Argumenten $y \in Y$ basiert ausschließlich auf der Existenz der Projektion und der Gültigkeit der Dreiecksungleichung in normierten Räumen, so dass der Unterschied in der Beweisführung zum Satz 3.13 lediglich darin liegt, dass die Projektion nicht notwendigerweise eindeutig ist.*

Satz 4.6. *Seien $y \in Y$ und $\rho \geq 0$ gegeben. Sei $\{y_n\}_{n \in \mathbb{N}}$ eine beliebige Folge in Y , so dass $y_n \rightarrow y$ für $n \rightarrow \infty$. Sei $\{u_n\}_{n \in \mathbb{N}} \subset M_\rho$ eine Menge der Repräsentanten von Lösungen der Minimierungsprobleme*

$$\min_{u \in M_\rho} \|F(u) - y_n\|_Y$$

4 Quasilösungen bei nichtlinearen Identifikationsproblemen

Dann ist jeder schwach*-Häufungspunkt der Menge $\{u_n\}_{n \in \mathbb{N}}$ ein Minimierer des Problems

$$\min_{u \in M_\rho} \|F(u) - y\|_Y$$

Ist der Minimierer eindeutig, so konvergiert die gesamte Folge $\{u_n\}_{n \in \mathbb{N}}$ schwach* gegen diesen Minimierer.

Beweis. Die Menge $\{u_n\} \subset M_\rho$ ist beschränkt, so dass mindestens eine schwach*-konvergente Teilfolge $\{u_k\}$ in M_ρ existiert. Sei durch \hat{u} der schwach* Grenzwert dieser Folge bezeichnet. Für die entsprechende Bildfolge $\{F(u_k)\}$ gilt wegen Annahme 4.1 $q_k \rightarrow F(\hat{u}) =: \hat{q}$ und damit

$$\begin{aligned} \|\hat{q} - y\| &\leq \liminf_{k \rightarrow \infty} \|q_k - y\| \\ &\leq \liminf_{k \rightarrow \infty} \|q_k - y_k\| + \|y_k - y\| \\ &= \liminf_{k \rightarrow \infty} \|q_k - y_k\| = \liminf_{k \rightarrow \infty} d(\rho, y_k) = d(\rho, y) \end{aligned}$$

mit $y_k \rightarrow y$ und $d(\rho, y_k) := \|q_k - y_k\| = \inf_{q \in Q_\rho} \|q - y_k\|$ und der Lipschitzstetigkeit der Distanzfunktion für jedes $\rho \geq 0$ (s. Beweis Lemma 4.4 und Bemerkung 4.5). Somit ist \hat{u} der Minimierer des Problems.

Gilt weiterhin, dass der Minimierer $\hat{u} \in M_\rho$ eindeutig ist, so konvergiert die gesamte Folge $\{u_n\}_{n \in \mathbb{N}}$ gegen diesen Minimierer und das Problem (4.1.3) ist stabil nach der Definition 3.5. \square

4.2 ZUR ERMITTLUNG DES REGULARISIERUNGSPARAMETERS

Mit Existenz und Stabilität der Quasilösungen wird die Erfüllung der ersten Forderung zur Definition 3.5(i) erreicht. Definition 3.5(ii) ist für alle $y \in R(F)$, bei der Forderung die entsprechende u^* -Minimum-Norm-Lösung sei eindeutig, offensichtlich ebenfalls erfüllt. Zur Ivanov-Regularisierung von nichtlinearen Operatorgleichungen bleibt also noch die Existenz einer Parameterwahlregel zu nennen, welche die schwach* Konvergenz der Quasilösungen zu der exakten Lösung sichert.

Im Kapitel 3 wurden die bekannten Strategien der Parameterwahl im Bezug auf die Ivanovs Methode im linearen Fall ausgiebig diskutiert. Dabei hat sich ergeben, dass für die theoretische Analyse die a-posteriori Strategien am wichtigsten sind. Deswegen wird in diesem Kapitel von Anfang an auf das Diskrepanzprinzip von Morozov gesetzt.

Als erstes sollte man sich vergewissern, dass diese Regel für jedes gegebene y^δ , δ und $\tau > 1$ die entsprechenden Parameter ρ liefert. Einen bereits bewährten Nachweisweg dazu würde die Stetigkeit der Distanzfunktion bieten s. Kapitel 3.3. Dafür ist der Satz 3.9 jedoch nicht direkt anwendbar, da der Beweis dort sich im Wesentlichen auf die Konvexität der Bildmenge Q_ρ bezieht. So muss für die nichtlineare Operatorgleichungen die Stetigkeit der Distanzfunktion aufs neue nachgewiesen werden. Dazu wird vorerst die Unterhalbstetigkeit gezeigt.

Lemma 4.7. Für jedes $y \in Y$ ist die Distanzfunktion $\rho \mapsto d(\rho, y)$ beschränkt, monoton fallend und unterhalb-stetig.

Beweis. Die Distanzfunktion $d(\rho, y)$ ist *beschränkt*, denn es gilt für jedes $\rho \geq 0$ und jedes $y \in Y$

$$(4.2.1) \quad d(\rho, y) = \inf_{u \in M_\rho} \|F(u) - y\| \leq \|F(u^*) - y\| < \infty$$

Dies schließt die wesentlichen Unstetigkeitsstellen der Funktion aus.

Die Distanzfunktion $d(\rho, y)$ ist *monoton fallend*, denn mit $M_{\rho_1} \subset M_{\rho_2}$ für $\rho_1 < \rho_2$ gilt der Definition nach

$$(4.2.2) \quad \inf_{u \in M_{\rho_1}} \|F(u) - y\| \leq \inf_{u \in M_{\rho_2}} \|F(u) - y\|$$

Sei $d(\rho) := d(\rho, y)$ nicht *unterhalb-stetig*, das heißt es existiert eine Parameterfolge $\rho_n \rightarrow \rho_0$, so dass

$$(4.2.3) \quad d(\rho_0) > \liminf_{n \rightarrow \infty} d(\rho_n)$$

Nach Satz 4.3 existiert zu jedem ρ_n mindestens eine Lösung. So kann jeweils ein Repräsentant von Lösungen zu jedem n-ten Parameter betrachtet werden, im Weiteren mit der selben Nummer $n \in \mathbb{N}$ bezeichnet.

Wegen der Beschränktheit der Folge $d(\rho_n)$ nach dem Satz von Bolzano-Weierstraß existiert eine Teilfolge $d(\rho_{n_m}) \subset d(\rho_n)$ mit der Eigenschaft

$$(4.2.4) \quad \lim_{m \rightarrow \infty} d(\rho_{n_m}) = \underline{d} = \liminf_{n \rightarrow \infty} d(\rho_n)$$

4 Quasilösungen bei nichtlinearen Identifikationsproblemen

wo \underline{d} der kleinste Häufungspunkt der beschränkten Folge reeller Distanzwerte ist.

Die entsprechende Bildfolge der Lösungen $\{q_m := F(u_m)\}$ ist beschränkt und enthält im reflexiven Banachraum Y nach dem Satz von Eberlein-Smuljan eine schwach konvergente Teilfolge $\{q_{m_l}\} \subset \{q_l\}$ mit der Eigenschaft $q_l \rightharpoonup q$.

Die entsprechende Urbildfolge $\{u_l := F^{-1}q_l\}$ ist beschränkt in L^∞ und enthält nach dem Satz von Banach-Alaoglu eine schwach* konvergente Teilfolge $\{u_k\} \subset \{u_l\}$ mit der Eigenschaft $u_k \rightharpoonup^* u$.

Der Operator $F : U \supseteq D(F) \rightarrow Y$ ist schwach*-schwach abgeschlossen. Damit gilt für $u_k \rightharpoonup^* u$ und $q_k \rightharpoonup q$ als Teilfolge von $\{q_l\}$, dass der Grenzwert $q = F(u)$ ist.

Für die Folge $\{u_k\}$ gilt nach der Definition der Quasilösungen $\|u_k - u^*\| \leq \rho_k$. Mit der schwach*-unterhalb Stetigkeit der $\|\cdot\|_{L^\infty}$ -Norm s., z.B. [1] und $u_k \rightharpoonup^* u$ folgt

$$(4.2.5) \quad \|u - u^*\| \leq \liminf_{k \rightarrow \infty} \|u_k - u^*\| \leq \liminf_{k \rightarrow \infty} \rho_k = \rho_0$$

Aber $d(\rho)$ ist monoton fallend, so dass gelten muss $d(\|u - u^*\|) \geq d(\rho_0)$ und damit folgende Ungleichung

$$(4.2.6) \quad \|q - y\| \geq d(\rho_0) > \underline{d} = \liminf_{k \rightarrow \infty} \|q_k - y\|$$

und mit $q_k \rightharpoonup q$ ein Widerspruch zu der schwachen Unterhalbstetigkeit der Norm in reflexiven Banachräumen s., z.B. [1], womit die Distanzfunktion unterhalb-stetig ist. \square

Um die Stetigkeit der Distanzfunktion nachzuweisen bleibt zu zeigen, dass die Distanzfunktion oberhalb-stetig in jedem ρ_0 ist, das heißt

$$(4.2.7) \quad \forall \varepsilon > 0 \exists \sigma > 0, \text{ so dass für jedes } \rho \in (\rho_0 - \sigma, \rho_0 + \sigma) \text{ gilt } d(\rho, y) < d(\rho_0, y) + \varepsilon$$

Lemma 4.8. Für jedes $y \in Y$ ist die Distanzfunktion $\rho \mapsto d(\rho, y)$ oberhalb-stetig.

Beweis. Seien $y \in Y$ und $\rho_0 \in (0, \infty)$ beliebig aber fest gewählt. Nach Satz 4.3 existiert mindestens eine Lösung

$$(4.2.8) \quad u_0 = \arg \min_{u \in M_{\rho_0}} \|F(u) - y\|$$

Im Weiteren werden zwei Fälle betrachtet. Fall (a): - für die Lösung gilt $\|u_0 - u^*\| = \rho_1 < \rho_0$, und Fall (b): - es gilt $\|u_0 - u^*\| = \rho_0$.

(a) Da u_0 Lösung ist, gilt nach der Definition der Distanzfunktion (3.3.1)

$$d(\rho_0, y) = d(\rho_1, y)$$

Sei nun $0 < \sigma = \rho_0 - \rho_1$ gewählt. Dann gilt für jedes $\varepsilon > 0$ und jedes $\rho \in (\rho_0 - \sigma, \rho_0 + \sigma)$ die Ungleichung

$$d(\rho, y) < d(\rho_0, y) + \varepsilon$$

da die Distanzfunktion monoton fallend ist. Somit ist die Distanzfunktion oberhalb-stetig in ρ_0 .

(b) Der Nachweis in diesem Fall wird auf der Stetigkeit des Operators aufgebaut. Nach der Voraussetzung ist $F : U \supseteq D(F) \rightarrow Y$ stetig, das heißt für ein $\varepsilon > 0$ existiert ein $\sigma > 0$, so dass mit $\|u - u_0\| < \sigma$ folgt $\|F(u) - F(u_0)\| < \varepsilon$. Als nächstes wird für ein beliebiges $\rho \in (\rho_0 - \sigma, \rho_0 + \sigma)$ folgendes Element $\tilde{u} \in M_\rho$ konstruiert

$$\tilde{u} - u^* = \frac{\rho}{\rho_0}(u_0 - u^*)$$

mit der Eigenschaft $\|u_0 - \tilde{u}\| < \sigma$, denn es gilt

$$\|\tilde{u} - u_0\| = \|\tilde{u} - u^* - (u_0 - u^*)\| = \left\| \frac{\rho}{\rho_0}(u_0 - u^*) - (u_0 - u^*) \right\| = \|u_0 - u^*\| \left| \frac{\rho}{\rho_0} - 1 \right| = |\rho - \rho_0| < \sigma$$

Nach Konstruktion ist $\|\tilde{u} - u^*\| = \rho$ und somit $\tilde{u} \in M_\rho$, so dass

$$d(\rho, y) = \inf_{u \in M_\rho} \|F(u) - y\| \leq \|F(\tilde{u}) - y\| \leq \|F(\tilde{u}) - F(u_0)\| + \|F(u_0) - y\| < \varepsilon + d(\rho_0, y)$$

womit die Distanzfunktion nach Definition (4.2.7) oberhalb-stetig in ρ_0 ist.

Da aber $\rho_0 \in (0, \infty)$ beliebig gewählt war, ist die Distanzfunktion auf dem gesamten Intervall oberhalb-stetig. \square

Bemerkung 4.9. Die Unterhalbstetigkeit in Lemma 4.7 wurde auf Grund der schwach*-schwach Abgeschlossenheit des Operators $F : U \supseteq D(F) \rightarrow Y$ nachgewiesen. Zum Nachweis der Oberhalbstetigkeit in Lemma 4.8 wurde darüber hinaus die Stetigkeit der Abbildung gebraucht, so dass mit der Annahme 4.1 folgende Behauptung gilt.

Satz 4.10. Für jedes $y \in Y$ ist die Distanzfunktion $\rho \mapsto d(\rho, y)$ stetig.

Beweis. Nach Lemma 4.7 und Lemma 4.8 folgt sofort die Behauptung. \square

Zu jedem $y^\delta \in Y$, $0 < \delta < \|y^\delta\|$, $\tau > 1$ garantiert die Stetigkeit der Distanzfunktion die Existenz solcher Parameter ρ und entsprechender Lösungen u_ρ^δ , welche die Bedingungen zum Diskrepanzprinzip erfüllen

$$(4.2.9) \quad \delta \leq \inf_{u \in M_\rho} \|F(u) - y^\delta\| \leq \tau \delta$$

Um die Anwendbarkeit der Quasilösungsmethode bei solcher Parameterwahl für ein nichtlineares Problem (4.1.3) zu begründen, wird im folgenden Satz gezeigt, dass bei einer Datenfolge $y^{\delta_n} \rightarrow y$ für $\delta_n \rightarrow 0$ die entsprechenden Lösungen schwach* zu u^\dagger konvergieren.

Satz 4.11. Sei $\rho_n := \rho(\delta_n, y^{\delta_n})$ gewählt nach Diskrepanzprinzip (3.4.3) für ein $\tau > 1$. Dann gilt $u_{\rho_n}^{\delta_n} \rightharpoonup^* u^\dagger$ für $n \rightarrow \infty$.

Beweis. Offensichtlich konvergiert die Bildfolge der Lösungen $F(u_{\rho_n}^{\delta_n})$ gegen die exakte Daten $y = F(u^\dagger)$ für $n \rightarrow \infty$, denn es gilt

$$\|F(u_{\rho_n}^{\delta_n}) - y\| \leq \|F(u_{\rho_n}^{\delta_n}) - y^{\delta_n}\| + \|y^{\delta_n} - y\| \leq (\tau + 1)\delta_n$$

4 Quasilösungen bei nichtlinearen Identifikationsproblemen

Zum Nachweis der Behauptung wird die Beschränktheit der Parameterfolge ρ_n benötigt. Nach Lemma 4.4 gilt

$$d(\rho^\dagger, y^{\delta_n}) = |d(\rho^\dagger, y^{\delta_n}) - d(\rho^\dagger, y)| \leq \|y^{\delta_n} - y\| \leq \delta_n \quad \text{für jedes } n \in \mathbb{N}$$

Weiterhin ist die Distanzfunktion monoton fallend (4.2.2) und es gilt

$$\delta_n \leq d(\rho_n, y^{\delta_n}) \leq \tau \delta_n \quad \text{für jedes } n \in \mathbb{N}$$

womit die Beschränktheit der Parameterfolge mit ρ^\dagger folgt

$$\rho_n \leq \rho^\dagger \quad \text{für jedes } n \in \mathbb{N}$$

Somit ist jede Urbildfolge $\{u_{\rho_n}^{\delta_n}\}$ beschränkt und es existiert eine schwach*-konvergente Teilfolge (bezeichnet wieder durch die selben Indizes) $u_{\rho_n}^{\delta_n} \rightharpoonup^* u \in M_{\rho^\dagger}$. In der Umgebung $B_\rho(u^*)$ existiert eine eindeutige Minimum-Norm-Lösung $u^\dagger = F^{-1}y$ zu den exakten Daten $y \in R(F) \subset Y$ nach Annahme 4.2. Die Abbildung ist schwach*-schwach abgeschlossen und für die Bildfolge gilt wegen der starken Konvergenz auch $F(u_{\rho_n}^{\delta_n}) \rightarrow F(u^\dagger)$. Das heißt, der Grenzwert der Teilfolge $\{u_{\rho_n}^{\delta_n}\}$ ist u^\dagger . Nach dem Teilfolgen–Teilfolgen Argument konvergiert die gesamte Lösungsfolge schwach* gegen u^\dagger . \square

Der obere Satz zeigt also, dass die Ivanov Methode mit der Wahl der Parameter nach dem Diskrepanzprinzip auch im nicht-linearen Fall eine konvergente Regularisierung bietet. Nun kann die Frage nach den Konvergenzraten diskutiert werden.

4.3 KONVERGENZRATEN

Im Kapitel 3.5 wurden die Konvergenzraten bei der theoretischen Untersuchung auf der Basis klassischer Quelledarstellung hergeleitet. Dazu wurde die Glattheit des Subgradienten vorausgesetzt. Das Testproblem in 3.6 hatte die in der Theorie geforderte Glattheit jedoch nicht. Trotz dieser Gegebenheit konvergiert die Bregmandistanz mit $O(\delta)$ s. (3.6.8); Tabelle 3.1. In der Arbeit [30] diente solch ein Verhalten der Lösungen eines nicht-glaten Testproblems, der Entwicklung neuer Technik zur Herleitung der Konvergenzraten. Anstatt der klassischen Quelledarstellung wurden die so genannten *variational source conditions* formuliert, welche eine Voraussetzung unmittelbar an die Bregmandistanz stellen. Dabei wurde die Glattheitsannahme abgeschwächt und ein neuartiges Ergebnis zu Konvergenzraten auf der Basis einer Variationsungleichung nachgewiesen. Solch eine Technik wurde erfolgreich angewandt und weiterentwickelt, insbesondere für nicht-lineare Operatorgleichungen. In [31] wurde ein Zusammenhang zwischen den klassischen und variationellen Quelledarstellung untersucht und unter Anderem gezeigt, dass in Hilberträumen die Gültigkeit der variationellen aus der Gültigkeit der klassischen folgt [31, Prop. 4.5]. Andererseits scheint die variationelle Quelledarstellung einen sehr viel breiteren Gültigkeitsbereich für verschiedenen Inverse Probleme zu haben. In [23] wurde für Banachräume X, Y gezeigt, dass im Falle einer schwach-schwach-abgeschlossenen Abbildung $F : X \rightarrow Y$ und eines Funktionals $J : X \rightarrow \mathbb{R}$ - konvex, eigentlich (proper) und schwach-unterhalbstetig, die variationelle Quelledarstellung stets existiert, wenn das beschränkte Domain $M = \{x : J(x) \leq c, c > 0\}$ schwach-folgenkompakt ist [23, thm. 5]. Im Allgemeinen, vom theoretischen Standpunkt her, ist dieses Ergebnis im vorliegenden Fall nicht direkt anwendbar, da die Menge $M_\rho \subset U$ nur schwach*-kompakt ist. Daher werden hier die Konvergenzraten basierend auf einer Annahme über Existenz der variationeller Quelledarstellung s. [38] untersucht.

Zur Herleitung der Konvergenzraten wird die Bregmandistanz

$$D_J^{\xi^\dagger}(u_\rho^\delta, u^\dagger) = J(u_\rho^\delta) - J(u^\dagger) - \langle \xi^\dagger, u_\rho^\delta - u^\dagger \rangle$$

betrachtet mit einem nicht leeren Subdifferential $\xi \in \partial J(u)$ für ein konvexes Funktional $J : U \rightarrow \mathbb{R}^+$.

Die variationelle Quelledarstellung wird nun wie folgt formuliert

Annahme 4.12. Für alle $u := \{u \in D(F) : J(u) \leq J(u^\dagger)\}$ existiert $\beta \in [0, 1)$, so dass gilt

$$(4.3.1) \quad -\langle \xi^\dagger, u - u^\dagger \rangle \leq \beta D_J^{\xi^\dagger}(u_\rho^\delta, u^\dagger) + \varphi(\|F(u) - F(u^\dagger)\|)$$

mit $\xi^\dagger \in \partial J(u^\dagger)$ und einer Indexfunktion $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ (d.h. φ - monoton wachsend und $\lim_{t \rightarrow 0} \varphi = 0$).

Satz 4.13. Sei die Annahme 4.12 erfüllt. Sei $J(u) := \|u\|_U$ und $\rho = \rho(\delta, y^\delta)$ nach Diskrepanzprinzip (3.4.3) gewählt. Dann gelten folgende Konvergenzraten

$$D_J^{\xi^\dagger}(u_\rho^\delta, u^\dagger) \leq \frac{1}{1-\beta} \varphi((\tau+1)\delta)$$

4 Quasilösungen bei nichtlinearen Identifikationsproblemen

Beweis. $J(u) := \|u\|_{L^\infty(\Omega)}$ ist konvex, so dass ein $\xi^\dagger \in \partial J(u^\dagger)$ existiert. Für alle Lösungen u_ρ^δ gewählt nach Diskrepanzprinzip ist die Bedingung $\|u_\rho^\delta\| \leq \|u^\dagger\|$ in Annahme 4.12 erfüllt und es gilt

$$\begin{aligned}
 D_J^{\xi^\dagger}(u_\rho^\delta, u^\dagger) &= \|u_\rho^\delta\| - \|u^\dagger\| - \langle \xi^\dagger, u_\rho^\delta - u^\dagger \rangle \\
 &\leq \beta D_J^{\xi^\dagger}(u_\rho^\delta, u^\dagger) + \varphi \|F(u_\rho^\delta) - F(u^\dagger)\| \\
 &\leq \beta D_J^{\xi^\dagger}(u_\rho^\delta, u^\dagger) + \varphi (\|F(u_\rho^\delta) - y^\delta\| - \|y^\delta - F(u^\dagger)\|) \\
 &\leq \beta D_J^{\xi^\dagger}(u_\rho^\delta, u^\dagger) + \varphi(\tau + 1)\delta
 \end{aligned}$$

□

In [38] ist der obere Satz in einer allgemeineren Form gezeigt.

4.4 ANWENDUNG: NUMERISCHE LÖSUNGEN, SEMI-LINEARER FALL

Wie auch im linearen Fall, wird hier zur Prüfung der Anwendbarkeit die Rekonstruktion der Temperaturquelle $u(x) \in U$ zum beobachteten Zustand $y \in Y$ betrachtet. Mit dem Unterschied, dass das Temperaturfeld durch einen nicht linearen Differentialoperator (semi-linear) beschrieben wird.

Der nichtlineare Operator $F : U \rightarrow Y$ wird durch ein Randwertproblem mit homogenen Neumann Randwertbedingungen

$$(4.4.1) \quad \begin{cases} -\Delta y + cy^3 = u & \text{in } \Omega, \\ \partial_\nu y = 0 & \text{auf } \partial\Omega, \end{cases}$$

für ein gegebenes $c > 0$ definiert.

Für die numerische Lösung wird analog zum linearen Fall ein Gleichungssystem für die semi-smooth Newton Methode aufgestellt. Wegen der nicht linearen Abbildung $F : U \rightarrow Y$ wird hier auf die Konstruktion des Gleichungssystems etwas detaillierter als im linearen Fall eingegangen, basierend auf Ideen und Algorithmen aus [15], [16], [17].

Die Restriktionen des Problems (4.1.3) werden über eine charakteristische Funktion $\delta_M : U \rightarrow \mathbb{R} \cup \{\infty\}$ mit

$$\delta_M(u) := \begin{cases} 0 & \text{wenn } u \in M_\rho \\ \infty & \text{wenn } u \in U \setminus M_\rho \end{cases}$$

aufgefasst und eine äquivalente Problemformulierung

$$(4.4.2) \quad \min_{u \in U} G(u) + \delta_M(u)$$

betrachtet. Das Funktional $G : U \rightarrow \mathbb{R}$ ist durch

$$G(u) := \frac{1}{2} \|F(u) - y^\delta\|_Y^2$$

gegeben.

Für $U = L^\infty(\Omega)$ und $Y = L^2(\Omega)$ besitzt das Randwertproblem (4.4.1) für jedes $u \in L^\infty(\Omega)$ genau eine schwache Lösung $y \in H^1(\Omega)$ die stetig auf $\bar{\Omega}$ ist s., [55, Satz 4.10]. Damit gilt für das Bild des Operators

$$R(F) \subset H^1(\Omega) \cap C(\bar{\Omega}) \subset L^2(\Omega)$$

Weiterhin ist F Fréchet-differenzierbar von $L^\infty(\Omega)$ in $H^1(\Omega) \cap C(\bar{\Omega})$ [55, Satz 4.17] und somit auch das Funktional $G(u)$ in (4.4.2). Für die Fréchet-Ableitung von $G : U \rightarrow \mathbb{R}$ an der Stelle (\bar{u}) in Richtung h gilt mit der Bezeichnung $\bar{y} := F(\bar{u})$

$$(4.4.3) \quad \begin{aligned} G'(\bar{u})h &= \langle \bar{y} - y^\delta, F'(\bar{u})h \rangle \\ &= \langle F'^*(\bar{u})(\bar{y} - y^\delta), h \rangle \end{aligned}$$

4 Quasilösungen bei nichtlinearen Identifikationsproblemen

So entsteht folgendes Randwertproblem zur Ermittlung von $\bar{p} := G'(\bar{u}) = F'(u)^*(\bar{y} - y^\delta)$ s., [55, Sec. 4.6]

$$(4.4.4) \quad \begin{cases} -\Delta p + 3cy^2 p = y - y^\delta & \text{in } \Omega \\ \partial_\nu p = 0 & \text{auf } \partial\Omega \end{cases}$$

Im nächsten Schritt soll die Notwendige Optimalitätsbedingung zum Problem (4.4.2) formuliert werden. Dabei wird der Beweis von [16, Prop. 2.2] zugrunde gelegt.

Satz 4.14. *Seien $\bar{u} \in M_\rho$ ein lokaler Minimierer des Problems (4.4.2) und $\bar{p} = G'(\bar{u})$. Dann gilt*

$$(4.4.5) \quad -\bar{p} \in \partial\delta_M(\bar{u})$$

Beweis. Mit Satz 4.3 ist die Quasilösung $\bar{u} \in M_\rho$ ein lokales Minimum von (4.4.2). Damit gilt für hinreichend kleine $t > 0$ und alle $u \in U$ die Ungleichung

$$(4.4.6) \quad G(\bar{u}) + \delta_M(\bar{u}) \leq G(\bar{u} + t(u - \bar{u})) + \delta_M(\bar{u} + t(u - \bar{u}))$$

Das Funktional $\delta_M(u)$ ist konvex, so dass folgt

$$\begin{aligned} \delta_M(\bar{u} + t(u - \bar{u})) &= \delta_M(tu + (1-t)\bar{u}) \\ &\leq t\delta_M(u) + \delta_M(\bar{u}) - t\delta_M(\bar{u}) \end{aligned}$$

und eingesetzt in (4.4.6)

$$\begin{aligned} 0 &\leq G(\bar{u} + t(u - \bar{u})) - G(\bar{u}) + \delta_M(\bar{u} + t(u - \bar{u})) - \delta_M(\bar{u}) \\ &\leq G(\bar{u} + t(u - \bar{u})) - G(\bar{u}) + t(\delta_M(u) - \delta_M(\bar{u})) \end{aligned}$$

Die Grenzwertbildung $t \searrow 0$ ergibt

$$(4.4.7) \quad 0 \leq \lim_{t \searrow 0} \frac{G(\bar{u} + t(u - \bar{u})) - G(\bar{u})}{t} + \delta_M(u) - \delta_M(\bar{u}) = \langle G'(\bar{u}), u - \bar{u} \rangle + \delta_M(u) - \delta_M(\bar{u})$$

Das Randwertproblem (4.4.4) ist bzgl. p linear und hat für jedes $\bar{y} \in L^2(\Omega)$ eine eindeutige Lösung $\bar{p} \in L^2(\Omega)$, so dass insbesondere $\bar{p} \in L^1(\Omega) \subset U^*$ ist.

Mit (4.4.7) ergibt sich

$$\delta_M(u) \geq \delta_M(\bar{u}) + \langle -\bar{p}, u - \bar{u} \rangle$$

also eine Ungleichung, welche für alle $u \in U$ gilt und der Definition nach einen Subgradienten $-\bar{p} \in \delta_M(\bar{u})$ liefert. \square

Im nächsten Schritt soll zum Aufbau des Algorithmus die Optimalitätsbedingung (4.4.5) in eine Funktionalgleichung transformiert werden. Für solch einen Schritt muss der Lösungsraum als Hilbertraum vorausgesetzt werden s. [32, Thm. 4.41]. Hier können die Funktionenräume in Betracht anschließender Diskretisierung wie folgt vorausgesetzt werden

$$U \equiv Y \equiv \mathbb{R}^N$$

so, dass die vorgenommene Transformation formal durchgeführt werden kann.

Durch die Multiplikation mit beliebigem $\gamma > 0$ und Addition von \bar{u} werden die Optimalitätsbedingungen transformiert in

$$\bar{u} - \gamma \bar{p} \in (Id + \gamma \partial \delta_M)(\bar{u})$$

Die Abbildung $(Id + \gamma \partial \delta_M)$ ist surjektiv, so dass bei einer eindeutigen Lösung \bar{u} folgende Gleichung erfüllt ist

$$\bar{u} = (Id + \gamma \partial \delta_M)^{-1}(\bar{u} - \gamma \bar{p})$$

Mit einem eigentlichen, konvexen und unterhalb-stetigen Funktional $\delta_M(u)$ kann in Betracht anschließender Diskretisierung der Funktionenräume $U \equiv Y \equiv \mathbb{R}^N$ beim weiteren Schritt Lemma 6.134 aus [7] angewendet werden, so dass mit der Bezeichnung $v := \bar{u} - \gamma \bar{p}$ für jedes $\gamma > 0$ punktweise fast überall gilt

$$(4.4.8) \quad \bar{u} = \arg \min_u \frac{1}{2\gamma} \|u - v\|_Y^2 + \delta_M(u) = \begin{cases} \rho & \text{wenn } v > \rho \\ v & \text{wenn } |v| \leq \rho \\ -\rho & \text{wenn } v < -\rho \end{cases}$$

Die Notwendige Optimalitätsbedingung bekommt somit die Form

$$\bar{u} = \text{proj}_{[-\rho, \rho]}(\bar{u} - \gamma \bar{p})$$

Nun kann ein nichtlineares Gleichungssystem aus zwei Differentialgleichungen mit homogenen Neumann Randwertbedingungen und einer funktionalen Gleichung, welche durch *aktive Mengen* $\mathcal{A}_+ := \{x \in \Omega : u(x) - \gamma p(x) > \rho\}$, $\mathcal{A}_- := \{x \in \Omega : u(x) - \gamma p(x) < -\rho\}$ und *inaktive Menge* $\mathcal{I} := \{x \in \Omega : |u(x) - \gamma p(x)| \leq \rho\}$ gegeben ist, aufgestellt werden

$$(4.4.9) \quad \begin{cases} -\Delta y + cy^3 = u \\ -\Delta p + 3cy^2 p = y - y^\delta \\ \mathbb{1}_{\mathcal{I}}(u - \gamma p) + \mathbb{1}_{\mathcal{A}_+} \rho - \mathbb{1}_{\mathcal{A}_-} \rho = u \end{cases}$$

wo $\mathbb{1}_{\mathcal{I}}$, $\mathbb{1}_{\mathcal{A}_+}$, $\mathbb{1}_{\mathcal{A}_-}$ Indikatorfunktionen sind

$$\mathbb{1}_{\mathcal{A}} = \begin{cases} 1 & \text{für } x \in \mathcal{A}_{+/-} \\ 0 & \text{sonst} \end{cases} \quad \mathbb{1}_{\mathcal{I}} = \begin{cases} 1 & \text{für } x \in \Omega \setminus \{\mathcal{A}_+ \cup \mathcal{A}_-\} \\ 0 & \text{sonst} \end{cases}$$

Damit ergibt sich für den einzelnen Newton-Schritt für die Variationen $(\delta y, \delta p, \delta u)$ das formale Gleichungssystem

$$(4.4.10) \quad \begin{cases} -\Delta \delta y + 3cy^2 \delta y - \delta u = \Delta y - cy^3 - u \\ -\Delta \delta p + 3cy^2 \delta p + (6cyp - 1) \delta y = \Delta p - 3cy^2 p + y - y^\delta \\ -\gamma \mathbb{1}_{\mathcal{I}} \delta p - \mathbb{1}_{\mathcal{A}_+ \cup \mathcal{A}_-} \delta u = u - \mathbb{1}_{\mathcal{I}}(u - \gamma p) - \mathbb{1}_{\mathcal{A}_+} \rho + \mathbb{1}_{\mathcal{A}_-} \rho \end{cases}$$

Unter Verwendung der Finiten-Elemente-Methode soll nun die diskrete Schreibweise des formal linearisierten Systems (4.4.10) mit

4 Quasilösungen bei nichtlinearen Identifikationsproblemen

$K_h \in \mathbb{R}^{N \times N}$ - Stiffness-Matrix

$M_h \in \mathbb{R}^{N \times N}$ - Mass-Matrix

$O \in \mathbb{R}^{N \times N}$ - Null-Matrix

$D_I \in \mathbb{R}^{N \times N}$, $D_{\mathcal{A}_{+/-}} \in \mathbb{R}^{N \times N}$ Diagonal-Matrizen mit den Einträgen, gegeben durch $\mathbb{1}_I$, $\mathbb{1}_{\mathcal{A}_+}$, $\mathbb{1}_{\mathcal{A}_-}$

dargestellt werden. Im Unterschied zum linearen Fall muss hier zusätzlich eine weitere, durch die Funktion $f(x)$ gewichtete $N \times N$ Mass-Matrix $M_h(f)$, für die Darstellung der Multiplikation der Art $\int_{\Omega} f \delta y dx$ definiert werden.

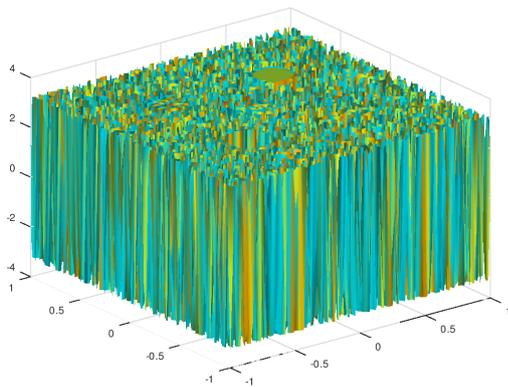
Für $y_h^{k+1} = y_h^k + \delta y$, $p_h^{k+1} = p_h^k + \delta p$, und $u_h^{k+1} = u_h^k + \delta u$ nimmt das System für einen Newton-Schritt folgende Gestalt an

$$(4.4.11) \quad B_k \begin{pmatrix} \delta y \\ \delta p \\ \delta u \end{pmatrix} := \begin{pmatrix} K_h + 3cM_h(y_h^k y_h^k) & O & -M_h \\ 6cM_h(y_h^k p_h^k) - M_h & K_h + 3cM_h(y_h^k y_h^k) & O \\ O & \gamma D_I & D_{\mathcal{A}_+} + D_{\mathcal{A}_-} \end{pmatrix} \begin{pmatrix} \delta y \\ \delta p \\ \delta u \end{pmatrix} \\ = - \begin{pmatrix} (K_h + cM_h(y_h^k y_h^k))y_h^k - M_h u_h^k \\ (K_h + 3cM_h(y_h^k y_h^k))p_h^k - M_h(y_h^k - y_h^\delta) \\ u_h^k - \mathbb{1}_I(u_h^k - \gamma p_h^k) - \mathbb{1}_{\mathcal{A}_+} \rho + \mathbb{1}_{\mathcal{A}_-} \rho \end{pmatrix} =: -b^k$$

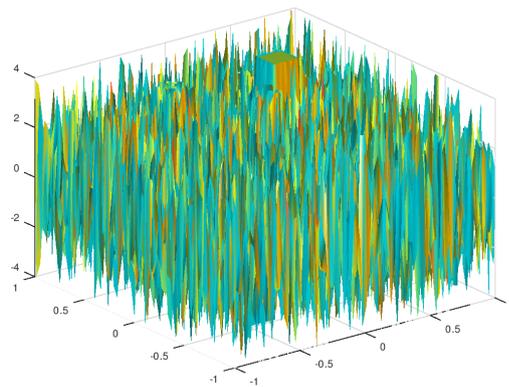
Im Unterschied zum linearen Fall weist dieses System einen stark erhöhten Rechenaufwand auf, da in jedem Newton-Schritt die Matrizen $M_h(y_h^k y_h^k)$ und $M_h(y_h^k p_h^k)$ aufs Neue ermittelt werden müssen. Dies führt zu der Notwendigkeit den Rechenaufwand bei der numerischen Umsetzung abzubauen, indem z.B. im Newton-Algorithmus auf die Dämpfung verzichtet wird. Die Bestimmung des optimalen Parameters gegeben im Algorithmus 3 wird ebenfalls vereinfacht, denn im linearen Fall wurde viel Aufmerksamkeit einem schnellen Umgang des Programms mit jedem beliebigen (unbequemem) Startwert ρ_0 geschenkt.

Die Abbildungen 3.2 zeigen die Lösungen zu $s = 0.1, 0.01, 0.001, 0.0001$ mit den entsprechenden Parametern $\rho = \rho(\delta, y^\delta) \approx 3.125, 3.819, 3.983, 3.998$ welche sich nach Diskrepanzprinzip ergeben. Im Unterschied zu den Ergebnissen in 3.2 sieht man hier deutlich, dass die Rekonstruktion später greift und erst im Bereich $0.01 < s \leq 0.01$ akzeptabel wird.

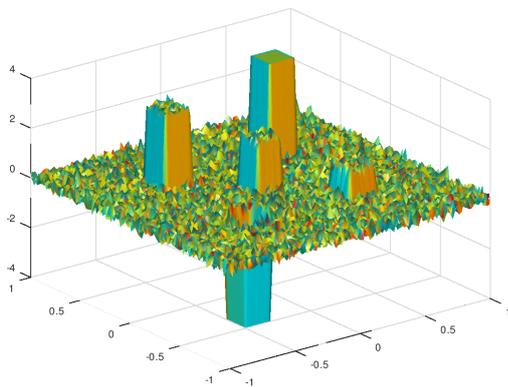
Dennoch zeigen die Ergebnisse die praktische Anwendbarkeit der Ivanov-Regularisierung für eine breite Klasse von nichtlinearen Problemen, insbesondere für den Fall, wenn der Lösungsraum nicht reflexiv ist. Dabei zeigt sich die Semi-Smooth-Newton Methode mit der lokal-superlinearen Konvergenz als ein sehr wirksames Framework, das darüber hinaus noch die Operator-Linearisierung innerhalb einzelner Schritte erlaubt, was zur einer schnelleren Rekonstruktion führt.



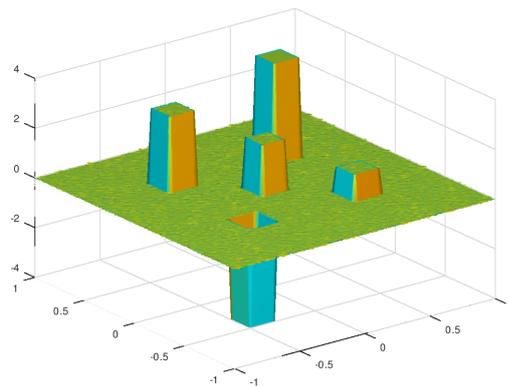
(a) u_ρ^δ für $s = 0.1\%$, $\rho \approx 3.125$



(b) u_ρ^δ für $s = 0.01\%$, $\rho \approx 3.819$



(c) u_ρ^δ für $s = 0.001\%$, $\rho \approx 3.983$



(d) u_ρ^δ für $s = 0.0001\%$, $\rho \approx 3.998$

Abbildung 4.1: Die Quasilösungen u_ρ^δ für unterschiedliche Störungslevel

Algorithm 3 Parameter choice for semi-linear case

Input: $y^\delta, \delta, \tau > 1, \rho_0 \geq 0$

- 1: set $y_h^k = 0, p_h^k = 0, u_h^k = 0$
- 2: **for** $k = 0, \dots$ **do**
- 3: compute y_h^k, p_h^k, u_h^k , conv using Algorithm 1 for $\rho_k, y_h^{k-1}, p_h^{k-1}, u_h^{k-1}$
- 4: compute $d_k = \|y_k - y^\delta\|_2$
- 5: **if** $d_k < \delta$ and conv = true **then**
- 6: **break**
- 7: set $\rho_{k+1} = \rho_k + \rho_0$
- 8: set $\rho_0 = \rho_k, \Delta\rho \leftarrow \Delta\rho/2, y_h^0 = y_h^k, p_h^0 = p_h^k, u_h^0 = u_h^k$
- 9: **for** $k = 0, \dots$ **do**
- 10: compute y_h^k, p_h^k, u_h^k , conv using Algorithm 1 for $\rho_k, y_h^{k-1}, p_h^{k-1}, u_h^{k-1}$
- 11: compute $d_k = \|y_k - y^\delta\|_Y$
- 12: **if** $\delta \leq d_k \leq \tau\delta$ and conv = true **then**
- 13: **break**
- 14: **else if** $d_k > \tau\delta$ or conv = false **then**
- 15: set $\Delta\rho \leftarrow \Delta\rho/2$
- 16: set $\rho_{k+1} = \rho_k + \Delta\rho$
- 17: **else**
- 18: set $\rho_{k+1} = \rho_k - \Delta\rho$

Output: $y_h^k, p_h^k, u_h^k, \rho_k$

5 ZUSAMMENFASSUNG UND AUSBLICK

In der vorliegenden Arbeit wurde die Quasilösungsmethode von Ivanov für den Fall einer schwach*-kompakten Lösungsmenge weiterentwickelt und die wirksame Anwendung dieser Methode gezeigt.

Dazu wurde zunächst die Existenz der Quasilösungen und die Stabilität (im Sinne der Definition 3.5) gezeigt, und im linearen Fall, bei zusätzlichen Glattheitsforderungen an den Datenraum, auch die Eindeutigkeit der Quasilösung nachgewiesen.

Die Untersuchung der Verfahrensweise der Regularisierungsmethode wurde auf Basis der Distanzfunktion (3.3.1) abgearbeitet. Die nachgewiesenen Eigenschaften der Distanzfunktion erlaubten es, die notwendigen und hinreichenden Bedingungen zu Parameteranforderungen aufzustellen. Darauf wurden die Parameterwahlstrategien untersucht und festgestellt, dass für diese Regularisierungsmethode im Allgemeinen keine reinen a-priori Strategien anwendbar sind. Dabei eignete sich die klassische a-posteriori Wahl nach dem Diskrepanzprinzip als die ideale Strategie für eine effiziente Umsetzung des Verfahrens. Mit dieser Parameterwahlregel wurde die schwach* Konvergenz des Verfahrens gezeigt.

Das Verhalten des Regularisierungsfehlers wurde mittels Bregmandistanz erfasst. Basierend auf der Annahme variationeller Quelldarstellung wurden die Konvergenzraten von $O(\delta)$ ermittelt.

Es soll vermerkt werden, dass im Falle eines gleichmäßig konvexen Lösungsraumes, die erhaltenen Resultate verstärkt werden können. Zum einen wird die schwach* zur schwachen Konvergenz. Und zum anderen garantiert die Methode die Norm-Konvergenz der Quasilösungen $\|u_\rho^\delta\|$ gegen $\|u^\dagger\|$ für $\delta \rightarrow 0$, so dass mit der Radon-Riesz-Eigenschaft sogar die starke Konvergenz vorliegt. Dies entspricht auch dem früher in [19] erhaltenem Resultat.

Um die praktische Anwendbarkeit der Ivanov-Regularisierung zu prüfen, ist das Verhalten der Quasilösungen an exemplarischen Testproblemen numerisch untersucht worden. Zur Ermittlung der Quasilösungen, vorbehaltlich der vorherigen Diskretisierung, ist der Einsatz der Semi-Smooth-Newton Methode begründet worden. Da die Quasilösungsmethode auf Hilberträumen im Gegensatz zu der Tichonov-Regularisierung keine glättende Eigenschaft besitzt, zeigt die Rekonstruktion ein erhöhtes Restrauschen. Dennoch demonstrieren die Ergebnisse, dass diese Rekonstruktion wirksam und brauchbar ist, insbesondere z.B. bei Problemen, wo das Hauptinteresse auf der Lokalisierung von Einflussfaktoren liegt.

Da die notwendigen und hinreichenden Bedingungen zur Parameterwahl (Kapitel 3.4), die Anwendung der a-priori und heuristischer Strategien prinzipiell, beim Vorhandensein zusätzlicher Informationen, nicht ausschließen, sind weitere Untersuchungen in diesem Bereich von großem Interesse.

Ein anderes interessantes Thema wäre der Aufbau iterativer Verfahren s. [39] zur Lösung nicht-linearer Inverser Probleme auf Basis der Ivanov-Regularisierung.

LITERATURVERZEICHNIS

- [1] H. W. Alt. *Lineare Funktionalanalysis*. Universitext. Springer, Berlin, Heidelberg, 2012. doi:10.1007/978-3-642-22261-0.
- [2] F. J. A. Artacho, J. M. Borwein, V. Martín-Márquez, and L. Yao. Applications of convex analysis within mathematics. *Math. Program.*, 148:49–88, 2014. doi:10.1007/s10107-013-0707-3.
- [3] A. Bakushinskii. Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion. *USSR Computational Mathematics and Mathematical Physics*, 24(4):181 – 182, 1984. doi:10.1016/0041-5553(84)90253-2.
- [4] F. Bauer and M. A. Lukas. Comparing parameter choice methods for regularization of ill-posed problems. *Mathematics and Computers in Simulation*, 81(9):1795 – 1841, 2011. doi:10.1016/j.matcom.2011.01.016.
- [5] H. Bauschke and P. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer International Publishing, 2018. doi:10.1007/978-3-319-48311-5.
- [6] J. M. Borwein and J. D. Vanderwerff. *Convex Functions: Constructions, Characterizations and Counterexamples*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2010. doi:10.1017/CB09781139087322.
- [7] K. Bredies and D. Lorenz. *Mathematische Bildverarbeitung: Einführung in Grundlagen und moderne Theorie*. Vieweg+Teubner, 2010. doi:10.1007/978-3-8348-9814-2.
- [8] K. Bredies and H. K. Pikkarainen. Inverse problems in spaces of measures. *ESAIM: Control, Optimisation and Calculus of Variations*, 19(1):190–218, 2013. doi:10.1051/cocv/2011205.
- [9] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer, New York, 2011. doi:10.1007/978-0-387-70914-7.
- [10] M. Burger and S. Osher. Convergence rates of convex variational regularization. *Inverse Problems*, 20(5):1411, 2004. doi:10.1088/0266-5611/20/5/005.
- [11] C. Clason. L^∞ fitting for inverse problems with uniform noise. *Inverse Problems*, 28(10):104007, 2012. doi:10.1088/0266-5611/28/10/104007.
- [12] C. Clason, K. Ito, and K. Kunisch. A convex analysis approach to optimal controls with switching structure for partial differential equations. *ESAIM: COCV*, 22(2):581–609, 2016. doi:10.1051/cocv/2015017.

- [13] C. Clason and B. Jin. A semismooth Newton method for nonlinear parameter identification problems with impulsive noise. *SIAM Journal on Imaging Sciences*, 5(2):505–536, 2012. doi:10.1137/110826187.
- [14] C. Clason and A. Klassen. Quasi-solution of linear inverse problems in non-reflexive Banach spaces. *Journal of Inverse and Ill-posed Problems*, 26(5):689–702, 2018. doi:10.1515/jiip-2018-0026.
- [15] C. Clason and K. Kunisch. Multi-bang control of elliptic systems. *Annales de l’Institut Henri Poincaré (C) Non Linear Analysis*, 31(6):1109 – 1130, 2014. doi:10.1016/j.anihpc.2013.08.005.
- [16] C. Clason and K. Kunisch. A convex analysis approach to multi-material topology optimization. *ESAIM: M2AN*, 50(6):1917–1936, 2016. doi:10.1051/m2an/2016012.
- [17] C. Clason and A. Schiela. Optimal control of elliptic equations with positive measures. *ESAIM: Control, Optimisation and Calculus of Variations*, 2015, 09 2015. doi:10.1051/cocv/2015046.
- [18] K. Deckelnick and M. Hinze. A note on the approximation of elliptic control problems with bang-bang controls. *Computational Optimization and Applications*, 51(2):931–939, 2012. doi:10.1007/s10589-010-9365-z.
- [19] I. N. Dombrovskaja and V. K. Ivanov. On the theory of certain linear equations in abstract spaces. *Sib. Math. Z*, 6(3):499–508, 1965.
- [20] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, Dordrecht, 1996. doi:10.1007/978-94-009-1740-8.
- [21] L. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.
- [22] J. Flemming. Theory and examples of variational regularization with non-metric fitting functionals. *J. Inverse Ill-Posed Probl.*, 18(6):677–699, 2010. doi:10.1515/JIIP.2010.031.
- [23] J. Flemming. Existence of variational source conditions for nonlinear inverse problems in Banach spaces. *Journal of Inverse and Ill-posed Problems*, 26, 11 2017. doi:10.1515/jiip-2017-0092.
- [24] J. Fletcher and W. B. Moors. Chebyshev sets. *Journal of the Australian Mathematical Society*, 98(2):161–231, 2015. doi:10.1017/S1446788714000561.
- [25] M. Grasmair. Generalized Bregman distances and convergence rates for non-convex regularization methods. *Inverse Problems*, 26(11):115014, 16, 2010. doi:10.1088/0266-5611/26/11/115014.
- [26] A. Göpfert, T. Riedrich, and C. Tammer. *Angewandte Funktionalanalysis*. Vieweg+Teubner, 2009. doi:10.1007/978-3-8348-9572-1.

- [27] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13(3):865–888, 2002. doi:10.1137/S1052623401383558.
- [28] M. Hinze. A variational discretization concept in control constrained optimization: The linear-quadratic case. *Computational Optimization and Applications*, 30:45–61, 2005. doi:10.1007/s10589-005-4559-5.
- [29] B. Hofmann. *Mathematik Inverser Probleme*. Teubner Verlagsgesellschaft, Stuttgart, 1999.
- [30] B. Hofmann, B. Kaltenbacher, C. Pöschl, and O. Scherzer. A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Problems*, 23(3):987–1010, 2007. doi:10.1088/0266-5611/23/3/009.
- [31] B. Hofmann and M. Yamamoto. On the interplay of source conditions and variational inequalities for nonlinear ill-posed problems. *Applicable Analysis*, 89:1705–1727, 11 2010. doi:10.1080/00036810903208148.
- [32] K. Ito and K. Kunisch. *Lagrange Multiplier Approach to Variational Problems and Applications*, volume 15 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008. doi:10.1137/1.9780898718614.
- [33] V. K. Ivanov. On linear problems which are not well-posed. *Dokl. Akad. Nauk SSSR*, 145(2):270–272, 1962.
- [34] V. K. Ivanov. On ill-posed problems. *Sib. Mat.*, 103(2):211–223, 1963.
- [35] V. K. Ivanov. On one class of ill-posed problems in vector topological spaces. *Sib. Mat.*, 4(4):832–839, 1965.
- [36] V. K. Ivanov. Ill-posed problems in topological spaces. *Sib. Mat.*, 10(5):1065–1074, 1969.
- [37] V. K. Ivanov, V. V. Vasin, and V. P. Tanana. *Theory of Linear Ill-Posed Problems and its Applications*. VSP, Utrecht, 2002.
- [38] B. Kaltenbacher and A. Klassen. On convergence and convergence rates for Ivanov and Morozov regularization and application to some parameter identification problems in elliptic PDEs. *Inverse Problems*, 34(5):055008, apr 2018. doi:10.1088/1361-6420/aab739.
- [39] B. Kaltenbacher, A. Klassen, and M. Previatti de Souza. On convergence and convergence rates for Ivanov and Morozov regularization and application to some parameter identification problems in elliptic PDEs. *Journal of Inverse and Ill-posed Problems*, 2019. doi:10.1515/jiip-2018-0093.
- [40] I. Leonard and K. Taylor. Essential supremum norm differentiability. *International Journal of Mathematics and Mathematical Sciences*, 8:433–439, 1985. doi:10.1155/S0161171285000473.

- [41] D. Lorenz and N. Worliczek. Necessary conditions for variational regularization schemes. *Inverse Problems*, 29(7):075016, 2013. doi:10.1088/0266-5611/29/7/075016.
- [42] S. Lu and P. Mathé. Heuristic parameter selection based on functional minimization: Optimality and model function approach. *Mathematics of Computation*, 82:1609–1630, 07 2013. doi:10.1090/S0025-5718-2013-02674-9.
- [43] V. A. Morozov. Choice of parameter in solving functional equations by the method of regularization. *Dokl. Akad. Nauk SSSR*, 175:1225–1228, 1967.
- [44] V. A. Morozov. *Methods for Solving Incorrectly Posed Problems*. Springer-Verlag, New York, 1984.
- [45] A. Neubauer and R. Ramlau. On convergence rates for quasi-solutions of ill-posed problems. *Electron. Trans. Numer. Anal*, 41:81–92, 2014. URL: <http://etna.math.kent.edu/volumes/2011-2020/vol41/abstract.php?vol=41&pages=81-92>.
- [46] T. Raus and U. Hämarik. *Heuristic Parameter Choice in Tikhonov Method from Minimizers of the Quasi-Optimality Function*. 2018. doi:10.1007/978-3-319-70824-9_12.
- [47] E. Resmerita. Regularization of ill-posed problems in Banach spaces: Convergence rates. *Inverse Problems*, 21:1303–1314, 08 2005. doi:10.1088/0266-5611/21/4/007.
- [48] A. Rieder. *Keine Probleme mit Inversen Problemen*. Vieweg+Teubner, 2003. doi:10.1007/978-3-322-80234-7.
- [49] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*. Springer, New York, 2009. doi:10.1007/978-0-387-69277-7.
- [50] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski. *Regularization Methods in Banach Spaces*. De Gruyter, Berlin, 2012. doi:10.1515/9783110255720.
- [51] T. I. Seidman and C. R. Vogel. Well posedness and convergence of some regularisation methods for non-linear ill posed problems. *Inverse Problems*, 5(2):227–238, 1989. doi:10.1088/0266-5611/5/2/008.
- [52] A. N. Tikhonov. On the regularization of ill-posed problems. *Dokl. Akad. Nauk SSSR*, 153(1):49–52, 1963.
- [53] A. N. Tikhonov. On the solution of ill-posed problems and the method of regularization. *Dokl. Akad. Nauk SSSR*, 151(3):501–504, 1963.
- [54] A. N. Tikhonov and V. Y. Arsenin. *Solution of Ill-Posed Problems*. V. H. Winston & Sons, Washington, D.C.: John Wiley & Sons, New York-Toronto, Ont.-London, 1977.
- [55] F. Tröltzsch. *Optimal Control of Partial Differential Equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Theory, methods and applications. doi:10.1090/gsm/112.

- [56] M. Ulbrich. *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, volume 11 of *MOS-SIAM Series on Optimization*. SIAM, Philadelphia, PA, 2011. doi:10.1137/1.9781611970692.
- [57] G. M. Vainikko and A. Y. Veretennikov. *Iteration Procedures in Ill-Posed Problems*. Nauka, Moscow, 1986.
- [58] V. Vasin, G. Skorik, E. Pimonov, and F. Kuchuk. New regularization algorithms for solving the deconvolution problem in well test data interpretation. *Scientific Research (Applied Mathematics)*, 1:387–399, 2010. doi:10.4236/am.2010.15051.
- [59] V. V. Vasin. The method of quasi-solutions by Ivanov is the effective method of solving ill-posed problems. *Journal of Inverse and Ill-posed Problems*, 16(6):537–552, 2008. doi:10.1515/JIIP.2008.028.
- [60] V. A. Vinokurov. A certain necessary condition for Tikhonov regularizability. *Dokl. Akad. Nauk SSSR*, 195(3):530–531, 1970.
- [61] D. Werner. *Funktionalanalysis*. Universitext. Springer, Berlin, Heidelberg, 2007. doi:10.1007/978-3-540-72536-7.
- [62] A. Yagola and V. Titarenko. Using a priori information about a solution of an ill-posed problem for constructing regularizing algorithms and their applications. *Inverse Probl. Sci. Eng.*, 15(1):3–17, 2007. doi:10.1080/17415970600573809.
- [63] A. G. Yagola. Error estimation for ill-posed problems with priori information. *Global Journal of Technology and Optimization*, 1:88–92, 2010.
- [64] A. G. Yagola, A. S. Leonov, and V. N. Titarenko. Data errors and an error estimation for ill-posed problems. *Inverse Problems in Engineering*, 10(2):117–129, 2002. doi:10.1080/10682760290031195.
- [65] A. G. Yagola and V. N. Titarenko. A posteriori error estimation for ill-posed problems on some sourcewise represented or compact sets. pages 425–442, 2002.
- [66] K. Yosida. *Functional analysis*. (Grundlehren der mathematischen Wissenschaften, Vol. 123). Springer - Berlin, Heidelberg, 1995 (Reprint of 1980). doi:10.1007/978-3-642-61859-8.
- [67] K. Yosida and E. Hewitt. Finitely additive measures. *American Mathematical Society*, 72(2):46–66, 1952. doi:10.1090/S0002-9947-1952-0045194-X.